On the Ordering of Communication Channels

by

Adithya Rajan

A Dissertation Presented in Partial Fulfillment
of the Requirement for the Degree
Doctor of Philosophy

Approved February 2014 by the
Graduate Supervisory Committee:

Cihan Tepedelenlioglu, Chair
Antonia Papandreou-Suppappola
Daniel Bliss
Oliver Kosut

ARIZONA STATE UNIVERSITY

May 2014

ABSTRACT

This dissertation introduces stochastic ordering of instantaneous channel powers of fading channels as a general method to compare the performance of a communication system over two different channels, even when a closed-form expression for the metric may not be available. Such a comparison is with respect to a variety of performance metrics such as error rates, outage probability and ergodic capacity, which share common mathematical properties such as monotonicity, convexity or complete monotonicity. Complete monotonicity of a metric, such as the symbol error rate, in conjunction with the stochastic Laplace transform order between two fading channels implies the ordering of the two channels with respect to the metric. While it has been established previously that certain modulation schemes have convex symbol error rates, there is no study of the complete monotonicity of the same, which helps in establishing stronger channel ordering results. Toward this goal, the current research proves for the first time, that all 1-dimensional and 2-dimensional modulations have completely monotone symbol error rates. Furthermore, it is shown that the frequently used parametric fading distributions for modeling line of sight exhibit a monotonicity in the line of sight parameter with respect to the Laplace transform order. While the Laplace transform order can also be used to order fading distributions based on the ergodic capacity, there exist several distributions which are not Laplace transform ordered, although they have ordered ergodic capacities. To address this gap, a new stochastic order called the ergodic capacity order has been proposed herein, which can be used to compare channels based on the ergodic capacity. Using stochastic orders, average performance of systems involving multiple random variables are compared over two different channels. These systems include diversity combining schemes, relay networks, and signal detection over fading channels with non-Gaussian additive noise. This research also addresses the problem of unifying fading distributions. This

unification is based on infinite divisibility, which subsumes almost all known fading distributions, and provides simplified expressions for performance metrics, in addition to enabling stochastic ordering.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

CHAPTER Page

LIST OF FIGURES

Chapter 1

INTRODUCTION

This age, rightly dubbed as the "information age" has witnessed tremendous and ever increasing growth in terms of communication technology and paraphernalia. Indeed, the world has become a highly connected and well networked place, and much of it is due to the foundations laid down by [1]. For example, Shannon theory forms the basis of the internet, which connects virtually every other person on the face of the earth; it triggered the development of deep space networks, permitting us to explore the heavens; terrestrial navigation using global positioning has become a possibility, thanks to digital communications; people can now connect while on the move as a result of development in the area of wireless communications.

Wireless systems such as cellular phones, wireless internet devices and bluetooth devices have become ubiquitous in the communication arena. Although diverse in their nature of operation, each of these devices can be thought of as a transmitter-receiver pair, where the transmission of the signal is over a wireless medium (channel). This transmission experiences random fluctuations in signal strength due to various physical factors that impair the wireless channel. This phenomenon is known as fading in wireless communication parlance, and stochastic models known as fading distributions are built to capture this effect in a system. Some typical fading distributions are summarized in Table 1.1.

For a wireless system, a number of performance metrics to assess the quality of the system have been defined in the communication theory literature. A few examples are (i) the *outage probability*, which gives a measure of how often the channel is not good enough to support reliable communication, (ii) the average symbol error rate

Name	**Probability Density Function**	**Parameters**
Rayleigh	$f(x) = \frac{x}{\sigma^2}\exp(-x^2/2\sigma^2)$	$\sigma \geq 0$
Rician	$f(x) = \frac{x}{\sigma^2}\exp(-\frac{x^2+a^2}{2\sigma^2})I_0(ax/\sigma^2)$	$a, \sigma \geq 0.$
Nakagami-m	$f(x) = \frac{2m^m}{\Gamma(m)\Omega^m}x^{2m-1}\exp(-mx^2/\Omega)$	$m \geq 0.5$, $\Omega > 0.$
Hoyt	$f(x) = \frac{1+q^2}{q\Omega}\exp\left(\frac{(1+q^2)^2x^2}{4q^2\Omega}\right)I_0\left(\frac{(1-q^4)x^2}{4q^2\Omega}\right)$	$q \in (0,1)$, $\Omega \geq 0.$
Weibull	$f(x) = (k/\lambda)(x/\lambda)^{k-1}\exp(x/\lambda)^k$	$k > 0$, $\lambda > 0.$
Log-normal	$f(x) = \frac{1}{x\sigma\sqrt{2\pi}}\exp(-\frac{(\log x-\mu)^2}{2\sigma^2})$	$\sigma \geq 0$, $\mu < \infty.$

Table 1.1: Commonly used fading distributions.

(SER), which is a measure of the average number of errors in signal transmission over all possible states of the fading channel, (iii) the ergodic capacity, which measures the maximum rate of signal transmission which can guarantee reliable communication, assuming that the communication duration is long enough to experience all possible channel states, and (iv) the diversity order, which is a measure of the number of independent communication channels used by the system. More precise definitions of these metrics will be provided in the forthcoming chapters.

Given the vast number of wireless systems with different purposes operating over fading channels, it is of interest to know how to decide whether one communication channel is superior to another. The performance of such systems are quantified by averaging a metric (e.g. bit or symbol error rates, or channel capacity) over the distribution of the random channel. Very often, when one channel is better than another in terms of a particular metric, it is also better with respect to another metric. However, this is not always true. Traditionally, answering this question has relied on single parameter comparisons between channels using characteristics such as diversity order, "amount of fading", Ricean factor and others [2]. These are parametric approaches that quantify how much fading the channel exhibits, but do not

provide a unified framework to compare channels across many different performance metrics. In this work, we propose to use stochastic orders to address this issue.

The theory of stochastic orders (or dominance) provides a comprehensive framework to compare two random variables (RVs) or vectors [3]. The simplest and most widely used stochastic order compares the cumulative distribution functions (CDF) of two RVs, which defines a partial order between pairs of RVs. When the RVs represent instantaneous signal to noise ratios (SNRs) in a fading environment, this corresponds to comparing the outage probabilities in a wireless communication context. Another stochastic order worth mentioning is the stochastic Laplace transform (LT) order, which compares the real-valued Laplace transforms of RVs. There are many other stochastic orders that capture comparisons of RVs in terms of size, and variability. Different than the related majorization theory [4, 5], which defines a partial order on deterministic vectors, stochastic orders apply to random variables. Stochastic ordering has become an indispensable tool in many increasingly diverse areas of probability and applied statistics over the past sixty years. Examples of such areas include reliability theory [6], actuarial sciences [7], risk analysis [8], economics [9] and stochastic processes [10]. However, the applications of this set of tools in physical layer wireless communications are surprisingly very few, although it has found numerous applications in communication networks (please see [10], [11], [3, Ch. 13-14] and the references therein).

One of the primary focus areas of this work is to employ the stochastic ordering theory to compare two wireless systems based on a particular set of metrics. In order to do so, the metric of choice, for e.g., the SER, needs to possess certain mathematical properties such as convexity or complete monotonicity. We recall that convex SERs have a negative first derivative and a positive second derivative with respect to the SNR. Since the convexity of the SER plays a critical role in various optimization

problems [12, 13], it has been the subject of investigation in the literature (see, for e.g., [14]). On the other hand, if all the successive derivatives of the SER also alternate in sign (referred to as complete monotonicity), then it is possible to express the SER as a positive mixture of decaying exponentials. This property proves to be extremely useful in comparing two fading channels using the LT order, as described in Chapter 3 of this dissertation. Despite the potential applications of complete monotonicity (c.m.) of the SER, it has not been investigated in its full generality previously. Consequently, we devote some attention to this issue in Chapter 4.

So far, we have mentioned how existing stochastic orders such as the LT order can be tailored to compare fading channels based on the average SER. Instead, if one wishes to compare fading channels based on the ergodic capacities, the LT order can still be used, however, not every pair of fading channels which can be compared based on the ergodic capacity is captured by the LT order. Further, it turns out that none of the existing stochastic orders completely characterizes every pair of fading channels ordered according to the ergodic capacity. We therefore develop a stochastic order, which is new to both information theory as well as stochastic ordering theory, which can be used to compare any two fading channels based on the ergodic capacity. Such comparisons can be made even in cases when a closed-form expression for the ergodic capacity is not analytically tractable.

While it is not difficult to verify whether two members of a given parametric fading distributions satisfy a particular stochastic order, there is no systematic method to do so for two distributions which do not belong to the same parametric fading family. This is partly due to the fact that there is no unified fading distribution which houses all of the known fading distributions and possesses properties which might be helpful from a stochastic ordering perspective. It is therefore of interest to see if the typical distributions used for multipath, shadowing and composite multipath/shadowing may

be unified under a common class with desirable analytical properties. Through a unification of fading distributions, it may be possible to obtain canonical expressions for the performance metrics of fading channels, thereby simplifying performance analysis of fading channel systems. The unified model may also permit the comparison of two different fading distributions with respect to system performance metrics such as the average SER, using stochastic ordering. It is therefore compelling to develop a unified study of fading models, with these goals in mind. The model proposed in Chapter 5.6 is based on nonnegative infinitely divisible (ID) distributions. A RV is said to be infinitely divisible, if it can be written as a sum of $n \geq 1$ independent and identically distributed (i.i.d.) RVs, for each n. Through this unification, it is observed that almost every distribution used to model multipath, shadowing and composite multipath/shadowing such as Rayleigh, Rician, Nakagami-m and lognormal are seen to be included in the class of ID RVs. The mathematical properties of this class of RVs makes it attractive from a performance analysis and stochastic ordering perspective.

1.1 Contributions and Organization of the Dissertation

In the previous section, we identified several gaps or shortcomings of techniques used for comparing fading channels using conventional methods. In doing so, we hinted at proposed solutions for the same. In what follows, we clearly identify and summarize the contributions of this work.

We propose to develop a systematic framework using the general theory of stochastic orders in Chapter 3, which can provide a means to compare two fading channels. In addition, we give a wide range of examples illustrating how different stochastic orders are appropriate for comparing systems using different metrics with analytical properties such as monotonicity, convexity, and complete monotonicity, which shed light into the connections between performance metrics such as error rates and er-

godic capacity. Additionally, we find the conditions under which a composite system formed using multiple fading links retains the order satisfied by the averages of performance metrics of individual systems. Such a system may involve combinations in parallel, in series, or otherwise, as may be seen in relay networks. Such a study permits the comparison of performance of systems, even in settings where closed-form expressions for the performance metrics are not tractable.

In Chapter 4, it is shown that the SER of an arbitrary multi-dimensional constellation in the absence of coding, when impaired by additive independent and identically distributed (i.i.d.) Gaussian noise under maximum likelihood (ML) detection can be represented as a product of a power of the SNR and a completely monotone function of SNR. This result also generalizes to SERs under compound Gaussian noise, which includes many non-Gaussian noise distributions such as Middleton class-A noise [15] and symmetric alpha-stable noise [16]. The SER of an arbitrary multi-dimensional constellation is shown to be completely monotone if the constellation matrix has a rank of one or two. Since complete monotonicity implies convexity, the SER is a convex function of the SNR, provided that the constellation matrix has a rank of one or two. For constellation matrices whose rank is greater than two, it is shown that the complete monotonicity of the SER depends on the constellation geometry and choice of prior probabilities. This work also describes a novel stochastic order for fading distributions, which can be used to order the average SERs of arbitrary multi-dimensional complex constellations over quasi-static fading channels, and generalizes the existing LT order on random variables.

A new stochastic order to compare two fading channels based on the ergodic capacity is introduced in Chapter 5. A detailed discussion of its properties, examples and extensions relevant to wireless communications, including the multiple-input multiple-output (MIMO) case is presented. Many parametric fading distribution fam-

ilies such as the Nakagami-m, Rician and Hoyt are observed to have the property that the ergodic capacity is monotone with respect to the line of sight (LoS) parameter for each of these distributions. Consequently, the instantaneous SNR of these fading channels serve as examples of ergodic capacity ordered random variables. The properties of this stochastic order are useful in obtaining comparisons of the performance of systems involving multiple SNR RVs. Such comparisons of ergodic capacities can be made even in cases when a closed-form expression for the ergodic capacity is not available. A MIMO extension of the definition of the ergodic capacity order, which can be used to order positive semi-definite symmetric random matrices is also given.

Chapter 6 proposes a new unified fading model using ID RVs. This unified model includes many popular fading models such as Rayleigh, Rician and Nakagami-m models, and any RV belonging to this class can be decomposed into the sum of $n \geq 1$ i.i.d. RVs, for each n. The properties of ID RVs are very relevant in asymptotic performance analysis of wireless systems and stochastic ordering of fading distributions within the class.

1.2 A Note on Notation

Here are some remarks on the notations used in this dissertation. The set of real numbers, positive integers and complex positive semi-definite symmetric matrices of size $n \times n$ are denoted by $\mathbb{R}$, $\mathcal{N}$, and $\mathbb{S}_+^n$ respectively, while all other sets are denoted using script font. For a finite set $\mathcal{B}$ the cardinality is denoted by $|\mathcal{B}|$, while the indicator function is defined as $I(x \in \mathcal{K}) = 1$, if $x \in \mathcal{K}$ and 0, otherwise. For a Radon measure $\mu(\cdot)$ on a Borel set of $\mathbb{R}^+$, $\mu(u)$ is used to represent $\mu([0, u])$. We write $f_1(x) = O(f_2(x))$, $x \to a$ to indicate that $\limsup_{x \to a}(f_1(x)/f_2(x)) < \infty$, and $f_1(x) \sim f_2(x)$, as $x \to a$ to indicate $\lim_{x \to a} f_1(x)/f_2(x) = 1$.

Vectors and matrices are denoted by boldface lower-case and upper-case letters

respectively. The transpose operator is denoted by $(\cdot)^{\mathrm{T}}$. For both vectors and matrices, $||\cdot||$ denotes the L^2 norm. The trace and determinant of a matrix $\mathbf{M}$ are denoted by $\mathbf{tr}\,\mathbf{M}$ and $\det(\mathbf{M})$ respectively. The j^{th} column vector and the k^{th} row vector of the matrix $\mathbf{M}$ are denoted by $\mathbf{m}^j$ and $\mathbf{m}_k$ respectively. The identity matrix is denoted by $\mathbf{I}$. If $a_i \in \mathbb{R}$, $i = 1, \ldots, N$, then $\mathbf{diag}(a_1, \ldots, a_N)$ is the diagonal matrix whose (i,i) element is $a_i, i = 1, \ldots, N$. The i^{th} smallest eigenvalue of $\mathbf{A} \in \mathbb{R}^{N \times N}$ is denoted by $\lambda_i(\mathbf{A})$, $i = 1, \ldots, N$. For a random variable X, $F_X(x)$ and $f_X(x)$ denote the cumulative distribution function (CDF) and the probability density function (PDF) respectively. $\mathbb{E}[g(X)]$ is used to denote the expectation of the function $g(\cdot)$ over the PDF of X. The multivariate real (circularly symmetric complex) Gaussian distribution with mean vector $\mathbf{a}$, and covariance matrix $\mathbf{C}$ is denoted by $\mathcal{N}(\mathbf{a}, \mathbf{C})$ ($\mathcal{CN}(\mathbf{a}, \mathbf{C})$). The Laplace transform and Laplace exponent of a nonnegative RV X are defined as $\phi_X(\cdot) = \mathbb{E}[\exp(-sX)]$ and $\gamma_X(s) = -\log \mathbb{E}[\exp(-sX)]$ respectively, which are defined for $s \geq 0$.

Chapter 2

MATHEMATICAL PRELIMINARIES

This chapter summarizes some of the mathematical concepts used in the forthcoming chapters. Notions such as complete monotonicity and stochastic ordering are used in almost all the chapters that follow, and as a result, these are discussed to reasonable depth here.

2.1 Some Special Classes of Functions

In what follows, some classes of functions such as completely monotone functions, Bernstein functions and Thorin-Bernstein functions are described. While the following discussion is sufficient for the applications in the scope of the current thesis, the interested reader is referred to [17] for a more comprehensive account.

2.1.1 Completely Monotone Functions

A function $g : (0, \infty) \to \mathbb{R}$ is completely monotone (c.m.), if and only if it has derivatives of all orders which satisfy

$$(-1)^n \frac{\mathtt{d}^n}{\mathtt{d}x^n} g(x) \geq 0, \tag{2.1}$$

for all $n \in \mathcal{N} \cup \{0\}$, where the derivative of order $n = 0$ is defined as $g(x)$ itself. The celebrated Bernstein's theorem [17] asserts that, $g : (0, \infty) \to \mathbb{R}$ is c.m. if and only if it can be written as a mixture of decaying exponentials:

$$g(x) = \int_0^\infty \exp(-ux) \mu(\mathtt{d}u), \tag{2.2}$$

which is a Lebesgue integral with respect to a positive measure μ on $[0, \infty)$. When $\mu(\mathtt{d}u) = \mu(u)\mathtt{d}u$, then we call μ as the *representing function* of g. It is straightforward to verify that c.m. functions are positive, decreasing and convex, and that positive linear combinations of c.m. functions are also c.m. The set of c.m. functions is denoted by $\mathcal{CM}$. $g(x) = 0.5\exp(-x)$ is an example of a c.m. function in a wireless communications context, as it corresponds to the SER of differential phase shift keying in AWGN.

A function $g : (0, \infty) \to \mathbb{R}$ is said to be completely monotone of order $\alpha \in \mathcal{N}$ if and only if $x^\alpha g(x)$ is c.m.. If g is c.m. of order α, then g is also c.m. of order β, where $0 \leq \beta < \alpha$. In [18, Theorem 1.3], it is shown that a necessary and sufficient condition for g to be c.m. of order α is that $g(x)$ must be represented in the form (2.2), where the integral converges for all $x > 0$. In addition, μ must be $\alpha - 1$ times differentiable, with the k^{th} derivative of $\mu(u)$ equal to zero at $u = 0$ for $0 \leq k \leq \alpha - 2$, and $\mathtt{d}^{\alpha-1}\mu(u)/\mathtt{d}u^{\alpha-1}$ non-negative, right-continuous and non decreasing.

2.1.2 Stieltjes Functions

The set of Stieltjes functions is a subclass of $\mathcal{CM}$, and is denoted by $\mathcal{S}$. A function $g : (0, \infty) \to [0, \infty)$ is said to belong to $\mathcal{S}$ if it admits the representation

$$g(x) = a/x + b + \int_{(0,\infty)} (x + u)^{-1}\mu(\mathtt{d}u) \, , \tag{2.3}$$

where $a, b \geq 0$, and μ is a nonnegative measure on $(0, \infty)$ which satisfies the convergence condition $\int_{(0,\infty)}(1 + u)^{-1}\mu(\mathtt{d}u) < \infty$. It is easy to show that any Stieltjes function is also a double Laplace transform of a nonnegative function. A necessary and sufficient condition for $x \mapsto g(x) \in \mathcal{S}$ is that $x \mapsto (g(x^{-1}))^{-1}$ also belongs to $\mathcal{S}$ [17, p. 66].

2.1.3 Bernstein Functions

A function $g : (0, \infty) \to \mathbb{R}$ is a Bernstein function, if $g(x) \geq 0, \forall x > 0$, and $\mathtt{d}g(x)/\mathtt{d}x$ is c.m. The set of all Bernstein functions is denoted by $\mathcal{BF}$. An equivalent representation of $g \in \mathcal{BF}$ is the following [17, p. 15]

$$g(x) = a + bx + \int_{(0,\infty)} (1 - \exp(-ux))\, \tau(\mathtt{d}u) \;, \tag{2.4}$$

for some $a, b \geq 0$, where τ is a nonnegative measure on $(0, \infty)$ satisfying $\int_{(0,1)} \tau(\mathtt{d}u) + \int_{[1,\infty)} u\tau(\mathtt{d}u) < \infty$. In this case, τ is known as the Levy measure of g. Bernstein functions are relevant to this work, because the instantaneous capacity function $g(x) = \log(1 + x)$ is a Bernstein function. Furthermore, the Laplace exponents of non-negative infinitely divisible RVs introduced in Chapter 5.6 are also Bernstein functions.

An important property is that the set $\mathcal{BF}$ is closed under positive linear combinations: if $g_i \in \mathcal{BF}$, and $a_i \geq 0$, $i = 1, \ldots, N$, then $\sum_{i=1}^{N} a_i g_i \in \mathcal{BF}$. Further, if $g_1(x)$ is c.m., and $g_2(x) \geq 0$ is a Bernstein function, then the composition $g_1(g_2(x))$ is completely monotonic. Some examples of Bernstein functions are $g(x) = x^\alpha$, for $0 < \alpha < 1$, $g(x) = x/(1+x)$ and $g(x) = \log(1+x)$. The representation of the capacity function $\log(1 + x)$ in the form (2.4) is known as Frullani's integral [19, p. 6], and is given by

$$\log(1 + x) = \int_0^\infty \left(1 - e^{-sx}\right) \frac{e^{-s}}{s} \mathtt{d}s \;. \tag{2.5}$$

2.1.4 Thorin-Bernstein Functions

A Bernstein function g is called a Thorin-Bernstein function [17, pp. 73-79], if it admits the representation given by (2.4), where $s\tau(s)$ is c.m.. The family of all Thorin-Bernstein functions is denoted by $\mathcal{TBF}$. A necessary and sufficient condition

for $g : (0, \infty) \to (0, \infty)$ to be in $\mathcal{TBF}$ is that g can be represented as follows [17, p. 73]:

$$g(x) = a + bx + \int_{(0,\infty)} \log\left(1 + x/u\right) \mu(\mathtt{d}u) \,, \tag{2.6}$$

for some $a, b \geq 0$ and μ is a positive measure on $(0, \infty)$, which satisfies the convergence condition $\int_0^1 |\log u| \mu(\mathtt{d}u) + \int_1^\infty u^{-1} \mu(\mathtt{d}u) < \infty$. In this case, μ is called the Thorin measure of g, and $\int_{(0,\infty)} \mu(\mathtt{d}u)$ is called the *mass* of the Thorin measure, or the *Thorin mass*.

We refer to any $g_2 \in \mathcal{TBF}$ which satisfies the property that $g_1(g_2(\cdot)) \in \mathcal{TBF}$ for all $g_1 \in \mathcal{TBF}$ as a *composable* Thorin-Bernstein function (we denote the set of all such functions by $\mathcal{CTBF}$). A necessary and sufficient condition for any g_2 to belong to $\mathcal{CTBF}$ is that $(\mathtt{d}g_2(x)/\mathtt{d}x)/g_2(x) \in \mathcal{S}$ [17, Theorem 8.4]. Functions belonging to the class $\mathcal{TBF}$ are of particular relevance to this work, since the Shannon capacity function $C(x) := \log(1 + x)$ not only belongs to $\mathcal{BF}$, but also belongs to $\mathcal{TBF}$, as seen from (2.5) and (2.6). Moreover, the Laplace exponents of generalized gamma convolutions described in Chapter 5.6 are also Thorin-Bernstein functions.

It is useful to define a multivariate extension of a Thorin-Bernstein function. A function $g : \mathbb{R}^m \to \mathbb{R}$ belongs to $\mathcal{TBF}_m$ if $g(x_1, \ldots, x_m)$ is a Thorin-Bernstein function in each argument, when all other arguments are treated as constants. Further, if g is composable in each variable when all other variables are fixed, then g is said to belong to the set $\mathcal{CTBF}_m$. An example of function in $\mathcal{CTBF}_m$ can be verified to be $g(x_1, \ldots, x_M) = \sum_{i=1}^M \alpha_i x_i, \alpha_i \geq 0, i = 1, \ldots, M$.

2.1.5 Matrix Functions

Let $\phi : \mathbb{R} \to \mathbb{R}$, and $\lambda_i \in \mathbb{R}$, $i = 1, \ldots, N$. If $\mathbf{D} = \mathbf{diag}\,(\lambda_1, \ldots, \lambda_N)$, we define $\phi(\mathbf{D}) = \mathbf{diag}\,(\phi(\lambda_1), \ldots, \phi(\lambda_N))$. If $\mathbf{A} \in \mathbb{S}_+^n$, so that

$$\mathbf{A} = \mathbf{U}\mathbf{diag}\,(\lambda_1(\mathbf{A}), \ldots, \lambda_N(\mathbf{A}))\,\mathbf{U}^{\mathrm{H}}\,, \tag{2.7}$$

where $\mathbf{U}$ is a unitary matrix, then we define $\phi(\mathbf{A}) = \mathbf{U}\phi(\mathbf{D})\mathbf{U}^{\mathrm{H}}$, provided ϕ is well defined on the eigenvalues of $\mathbf{A}$. In this way, $\phi(\mathbf{A})$ can be defined for all Hermitian matrices of any order [20]. In this work, the scalar function and its matrix extension are denoted using the same symbol, and the argument of the function defines the specific context. Matrix functions find applications in Section 5.4. We also make use of multivariate functions with matrix arguments in this work, which are defined through the Cauchy integral formula as given in [21]. While we refrain from providing the explicit definition here due to its rather technical nature, it suffices to note that such functions satisfy the following two properties [21], which will be used in our work.

Lemma 2.1.1. *If $\mathbf{A}_m \in \mathbb{S}_+^n$, $m = 1, \ldots, M$ then*

$$\mathbf{tr}\; f(\mathbf{A}_1, \ldots, \mathbf{A}_M) = \sum_{i_1=1}^{n} \cdots \sum_{i_M=1}^{n} f(\lambda_{i_1}(\mathbf{A}_1), \ldots, \lambda_{i_M}(\mathbf{A}_M))\,. \tag{2.8}$$

Lemma 2.1.2. *[21, Theorem 3.4], [21, p. 13] Let $\mathbf{A}_m \in \mathbb{S}_+^n$, $m = 1, \ldots, M$. If f is a multivariate matrix function well defined on the eigenvalues of $\mathbf{A}_m$, and ϕ is a univariate matrix function which is well defined on the eigenvalues of $f(\mathbf{A}_1, \ldots, \mathbf{A}_m)$, then $\phi(f(\mathbf{A}_1, \ldots, \mathbf{A}_m)) = (\phi \circ f)(\mathbf{A}_1, \ldots, \mathbf{A}_m)$.*

2.2 Transforms of Random Variables

In what follows, some transforms of densities of RVs are discussed. While this material is fairly standard, it is maintained here for the sake of completeness.

2.2.1 Laplace Transform

The (real-valued) Laplace transform (LT) of a nonnegative random variable X is defined as

$$\phi_X(\rho) = \mathbb{E}\left[\exp(-\rho X)\right], \ \rho \geq 0 \, . \tag{2.9}$$

Clearly, $\phi_X(\rho)$ can be obtained from the moment generating function (MGF) of X by evaluating the MGF at $-\rho$, and from the characteristic function evaluated at $j\rho$, where $j = \sqrt{-1}$. Due to the existence of the characteristic function of a RV, it follows that the LT of a nonnegative RV always exists, even if the density does not.

2.2.2 Shannon Transform

The Shannon transform of a nonnegative random variable X is defined as [22, pp. 44]:

$$\overline{C}^{(X)}(\rho) := \mathbb{E}\left[\log(1+\rho X)\right] \ , \rho \geq 0. \tag{2.10}$$

Two new representations of $\overline{C}^{(X)}(\rho)$, which are useful in this work are now obtained. Using (2.5), it is easy to show that (2.10) can be represented as a Laplace transform, given by

$$\overline{C}^{(X)}(\rho) = \int_0^\infty \exp(-u/\rho)\frac{1-\phi_X(u)}{u}\mathtt{d}u \, , \tag{2.11}$$

for $\rho > 0$. Using (2.2) with (2.11), it is immediate that $\overline{C}^{(X)}(\rho)$ is a c.m. function of $1/\rho$. A second representation of $\overline{C}^{(X)}(\rho)$ which can be derived from (2.11) shows that $\overline{C}^{(X)}(\rho)$ is also the Stieltjes transform [23, p. 325] of the complimentary CDF of X, when evaluated at $1/\rho$:

$$\overline{C}^{(X)}(\rho) = \int_0^\infty \frac{1-F_X(u)}{(1/\rho+u)}\mathtt{d}u \, , \tag{2.12}$$

where $\rho > 0$. Representation (2.12) is used in proving some properties of the ergodic capacity order discussed in Section 5.2.2. Additionally, (2.12) permits us to comment on the existence of $\overline{C}^{(X)}(\rho)$, which to the best of our knowledge has not been found elsewhere in the literature.

Proposition 2.2.1. *If $\overline{C}^{(X)}(\rho)$ exists for any $\rho \in (0, \infty)$, then $\overline{C}^{(X)}(\rho)$ exists for every $\rho \in (0, \infty)$.*

Proof. From (2.12), it is seen that $\overline{C}^{(X)}(\rho)$ is the Stieltjes transform of a real valued function. If the Stieltjes transform of a function exists at any point on $\mathbb{R}^+$, then it exists at all points on $\mathbb{R}^+$ [23, p. 326]. This completes the proof. □

We now provide examples of random variables for which the ergodic capacity is finite for $\rho < \infty$ using the following proposition:

Proposition 2.2.2. *Let $F_X(\cdot)$ denote the cumulative distribution function of a RV X. If for some $\delta \in (0, 1]$, $\int_0^t 1 - F_X(u)\,\mathtt{d}u = O(t^{1-\delta}), t \to \infty$, then $\overline{C}^{(X)}(\rho) < \infty$.*

Proof. First, observe that $\int_0^\infty (s+t)^{-1} \mathtt{d}\alpha(t)$ exists if $\alpha(t) = O(t^{1-\delta}), t \to \infty$, for some $\delta > 0$ [23, p. 330 (Theorem 3b)]. The proposition then follows by letting $\alpha(t) = \int_0^t 1 - F_X(u)\,\mathtt{d}u$. This completes the proof. □

In Proposition 2.2.2, the case of $\delta = 1$ is equivalent to the condition that the mean of X is finite. It is therefore straightforward to see that the ergodic capacity of fading distributions such as Nakagami-m and Rician is finite at all finite SNR, since these distributions have finite average power.

2.3 Stochastic Ordering

Stochastic orders are binary relations between random variables, which can be used to compare them based on a variety of criteria. The literature on stochastic

ordering, primarily in reliability theory and statistics, delineates numerous stochastic orders, many of which fall under the subclass of *integral stochastic orders*. We begin with a short description of the theory of integral stochastic orders, which can be found in [7, 3].

Let $\mathcal{G}$ denote a class of real valued functions $g : \mathbb{R}^+ \to \mathbb{R}$, and X and Y be RVs with CDFs $F_X(\cdot)$ and $F_Y(\cdot)$ respectively. We define the integral stochastic order with respect to $\mathcal{G}$ as [7]:

$$X \leq_{\mathcal{G}} Y \iff \mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)] \ , \ \forall g \in \mathcal{G}. \tag{2.13}$$

In this case, $\mathcal{G}$ is known as a generator of the order $\leq_{\mathcal{G}}$. A stochastic order can have more than one generator. For a given stochastic order, it is of interest to identify "large" generators which are useful in identifying the equivalence of two orders. The largest generator set of functions for a stochastic order contains all other generators, and is termed the *maximal generator* [7]. It is also of interest to find "small" generators which specify necessary conditions for the ordering of two RVs. We now give three examples of integral stochastic orders by specifying the corresponding generator set of functions $\mathcal{G}$.

2.3.1 Usual Stochastic Order

The usual stochastic order compares the magnitude of two RVs. In this case a small generator $\mathcal{G}$ is the set of all non-decreasing *indicator functions*: $\mathcal{G} = \{g(x) : g(x) = I[x > \rho], \rho \in \mathbb{R}\}$. From (2.13) it follows that this order is equivalent to comparing the CDFs of the RVs. Formally, we write

$$X \leq_{\text{st}} Y \iff F_X(x) \geq F_Y(x) \ \forall x \ . \tag{2.14}$$

To see the interpretation of this in the context of wireless channels, consider two channels to be compared, with channel powers $X := |h^X|^2$ and $Y := |h^Y|^2$, where

h^X and h^Y correspond to the complex channel gains of two wireless systems. The usual stochastic ordering of X and Y is equivalent to comparing their corresponding outage probabilities for all outage thresholds. The maximal generator for the usual stochastic order is the set of all increasing functions [7]. As a result, with the choice $g(x) = x$ in (2.13), we obtain $\mathbb{E}[X] \leq \mathbb{E}[Y]$ whenever $X \leq_{\text{Lt}} Y$, which agrees with the intuition that a larger random variable must have a larger mean value.

2.3.2 Convex Order

In this case $\mathcal{G}$ is the set of all convex functions, and the order is denoted as $X \leq_{\text{cx}} Y$. Since $g(x) = x$ and $g(x) = -x$ are both convex, from (2.13), we have $\mathbb{E}[X] = \mathbb{E}[Y]$ whenever X and Y are convex ordered. Therefore, convex ordering establishes that the RVs have the same mean and X is "less variable" than Y. Clearly, in the fading context, this can be used to identify channels with "less fading". Since $\leq_{\text{cx}}$ is a measure of variability, one would expect that a degenerate RV is less in the convex sense than any other RV with the same mean. Indeed, this is the case: If $F_X(x) = I[x \geq \mu]$, then $X \leq_{\text{cx}} Y$ for all RVs Y with $\mathbb{E}[Y] = \mu$. So the degenerate RV has an absolute minimum dispersion, as measured by the convex order, which is a consequence of Jensen's inequality.

Many performance metrics, such as channel capacity, error rates for different modulations [14] and coding schemes in wireless systems are convex or concave functions of the instantaneous SNR. Therefore, establishing convex ordering of two RVs can help us qualitatively measure the relative average performance of the corresponding systems. Note that if instead of convex functions, the class $\mathcal{G}$ is chosen as the set of all concave functions, one would get the same order with a reversal in the inequality.

Verifying the usual stochastic ordering of two RVs is straightforward through (2.14). What follows are easily testable sufficient conditions for $X \leq_{\text{cx}} Y$. Let

$S^{-}\left(g(x)\right)$ denote the number of sign changes of $g(x)$ as x increases over $[0,\infty)$, then $X \leq_{\text{cx}} Y$ if $\mathbb{E}\left[X\right] = \mathbb{E}\left[Y\right]$ and any of the following hold [3]:

$$S^{-}\left(f_Y\left(x\right) - f_X\left(x\right)\right) = 2 \text{ and the sign sequence is } +,-,+. \tag{2.15}$$

$$S^{-}\left(F_Y\left(x\right) - F_X\left(x\right)\right) = 1 \text{ and the sign sequence is } +,- \ , \tag{2.16}$$

where $f_X\left(\cdot\right)$ and $f_Y\left(\cdot\right)$ are the probability density functions (PDFs) of X and Y respectively.

Interestingly, to the best of our knowledge, although convex ordering of RVs is widely used in many other areas, it has never been used in physical-layer wireless communications.

Increasing Concave Order

Let X and Y be two RVs. Then X is said to be increasing concave ordered with Y (denoted $X \leq_{\text{icv}} Y$) if and only if $\mathbb{E}\left[g(X)\right] \leq \mathbb{E}\left[g(Y)\right]$ for all non-decreasing concave functions g, such that the expectations exist. A necessary and sufficient condition for the $\leq_{\text{icv}}$ order to hold [3] is the following:

$$X \leq_{\text{icv}} Y \iff \int_{-\infty}^{x} F_X\left(u\right) \mathtt{d}u \geq \int_{-\infty}^{x} F_Y\left(u\right) \mathtt{d}u \ , \forall x \ . \tag{2.17}$$

It is easy to see that by choosing $g(x) = x$, which is non-decreasing and concave, we obtain

$$X \leq_{\text{icv}} Y \implies \mathbb{E}\left[X\right] \leq \mathbb{E}\left[Y\right] \ , \tag{2.18}$$

whenever the expectations exist.

2.3.3 Laplace Transform Order

Similar to $\leq_{\text{st}}$ and $\leq_{\text{cx}}$, it is possible to order random variables based on LTs. In this case, $\mathcal{G} = \{g(x) : g(x) = -\exp(-\rho x),\ \rho \geq 0\}$, so that

$$X \leq_{\text{Lt}} Y \Longleftrightarrow \mathbb{E}\left[\exp(-Y\rho)\right] \leq \mathbb{E}\left[\exp(-X\rho)\right],\ \forall\ \rho > 0\ . \tag{2.19}$$

Interpreting $\exp(-\rho x)$ as being proportional to the instantaneous error rate (as in the case for differential-PSK (DPSK) modulation and Chernoff bounds for other modulations), LT ordering of the instantaneous SNRs in (2.19) can be interpreted as an inequality in the average error rates satisfied at all values of SNR ρ. One of the powerful consequences of LT ordering is that

$$X \leq_{\text{Lt}} Y \Longleftrightarrow \mathbb{E}\left[g(X)\right] \geq \mathbb{E}\left[g(Y)\right]\ , \tag{2.20}$$

for all c.m. functions $g(\cdot)$ [3, pp. 96]. A similar result with a reversal in the inequality states that

$$X \leq_{\text{Lt}} Y \Longleftrightarrow \mathbb{E}\left[g(X)\right] \leq \mathbb{E}\left[g(Y)\right]\ , \tag{2.21}$$

for all Bernstein functions $g(\cdot)$.

It is useful to mention that for any two RVs X and Y, $X \leq_{\text{cx}} Y \Rightarrow Y \leq_{\text{Lt}} X$, which follows from the fact that $-\exp(-\rho x)$ is concave in x for any $\rho > 0$. Hence, convex ordering provides a method to generate or verify LT ordering between two RVs. Indeed either of the conditions (2.15) or (2.16) together with equal mean values for X and Y imply that $X \leq_{\text{Lt}} Y$. Further, observe that $X \leq_{\text{st}} Y \Rightarrow X \leq_{\text{Lt}} Y$, which follows since $-\exp(-\rho x)$ is increasing in x for $\rho > 0$.

Another useful result for Laplace transform ordered RVs is the following:

Lemma 2.3.1. *Let $X_1, \ldots, X_M$ be independent and $Y_1, \ldots, Y_M$ be independent. If $X_m \leq_{Lt} Y_m$, $m = 1, \ldots, M$, then $g(X_1, \ldots, X_M) \leq_{Lt} g(Y_1, \ldots, Y_M)$, if $g(x_1, \ldots, x_M)$ is a c.m.d function of x_i, when all other arguments of g are viewed as constants.*

In a wireless communications context, let $\rho \geq 0$ be the average SNR, and ρX, ρY represent the instantaneous SNRs of two fading distributions. If $g(x)$ corresponds to the instantaneous symbol error rate $P_{\mathrm{e}}(\rho x)$ of a modulation scheme with c.m. error rate function, then (2.20) can be used to obtain comparisons of averages of symbol error rates over pairs of fading channels, even in cases where a closed-form expression for the same is intractable.

The definition of the LT order can be generalized to apply to nonnegative measure on $\mathbb{R}^+$. Let μ be a nonnegative measure. Its Laplace transform is defined as $\int_0^\infty \exp(-su)\mu(\mathtt{d}u)$. This is used to define the LT order between two nonnegative measures. If μ_1 and μ_2 are two nonnegative measures, then μ_1 is said to be dominated by μ_2 in the Laplace transform sense (denoted by $\mu_1 \leq_{\mathrm{Lt}} \mu_2$) whenever:

$$\int_0^\infty \exp(-su)\mu_1(\mathtt{d}u) \geq \int_0^\infty \exp(-su)\mu_2(\mathtt{d}u),\ \forall s \geq 0\ . \tag{2.22}$$

2.4 Introduction to Polyhedra and Polytopes

A set $\mathcal{P} \subseteq \mathbb{R}^N$ is a polyhedron if it is the intersection of finitely many closed half-spaces, i.e., $\mathcal{P} := \{\mathbf{x} | \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$, for some $\mathbf{A} \in \mathbb{R}^{M \times N}$, $\mathbf{b} \in \mathbb{R}^N$, and the inequality is applied component-wise [24]. A bounded polyhedron is referred to as a polytope. In a digital communications context, it is known that the decision region of a multi-dimensional constellation impaired by white Gaussian noise is a polyhedron.

A face of a polyhedron $\mathcal{P}$ is the intersection of $\mathcal{P}$ with a supporting hyperplane of $\mathcal{P}$. The dimension of a face is defined as the dimension of its affine hull. Faces of $\mathcal{P}$ of dimension zero, dimension one, and dimension $N-1$ are known as vertices, edges and facets of $\mathcal{P}$, respectively.

Some examples of polyhedra relevant to this research are described next. A polyhedral cone is a polyhedron, which is defined as $\mathrm{cone}(\mathcal{Y}) := \{\sum_{i=1}^M \lambda_i \mathbf{y}_i \mid \mathbf{y}_i \in \mathcal{Y}, \lambda_i \geq 0\}$, where $\mathcal{Y}$ is a non-empty set of points in $\mathbb{R}^N$. If the elements of $\mathcal{Y}$ are linearly inde-

pendent, then the polyhedral cone is called a simplicial cone. A well known result in combinatorial geometry literature is that any polyhedral cone admits a decomposition into simplicial cones [25, Lemma 1.40]. In this context, a decomposition of a polyhedron $\mathcal{P}$ is defined as a collection of sets $\{\mathcal{P}_1, \ldots, \mathcal{P}_R\}$, such that $\cup_{r=1}^{R} \mathcal{P}_r = \mathcal{P}$, and the intersection of any two sets in the decomposition is a common face of both, or the empty set.

2.5 Regular Variation

Some concepts from regular variation theory are now briefly touched upon, as this will be used for asymptotic analysis of performance metrics of fading channels. In vague terms, regular variation captures asymptotic polynomial-like behavior.

A real valued function $H(x) : \mathbb{R}^+ \mapsto \mathbb{R}^+$ is said to be regularly varying at ∞ (or at 0) if $\lim_{x\to\infty} H(tx)/H(x)$ (or $\lim_{x\to 0} H(tx)/H(x)$) exists, and is equal to t^r, with $0 < |r| < \infty$. It is said to be slowly varying at ∞ (or at 0) if $r = 0$. If H is differentiable with $h(x) = \mathtt{d}H(x)/\mathtt{d}x$, and $h(x) = mx^{m-1}l(x)$, where l is slowly varying at 0, then $H(x) \sim x^m l(x)$, as $x \to 0$.

Chapter 3

STOCHASTIC ORDERS FOR FADING CHANNELS

3.1 Motivation and Literature Survey

As discussed in Chapter 1, stochastic ordering theory has the potential to provide comparisons of wireless communication systems based on a variety of performance metrics, such as the error rate and the ergodic capacity. Despite its immediate applicability in wireless communications, to the best of our knowledge, there is no systematic exploitation of the general stochastic ordering theory which can be used to provide a means for comparing wireless systems. We now review the limited literature on the applications of stochastic orders in physical layer communications. Bounds on the per cell sum rate under arbitrary fading in the high SNR regime have been obtained using the aforementioned outage-based "usual stochastic order" in [26]. Stochastic ordering has also been applied to obtain bounds on the outage probability in Bluetooth piconets under Ricean fading in [27]. In [28], the usual stochastic order is used to bound monotone performance metrics in Ricean fading environments with beam selection. Reference [29] shows that stochastic ordering of the SNR between the sender and any two receivers is sufficient for the existence of a degraded channel in a layered erasure broadcast channel modelled using the binary expansion model.

All the above references use the usual stochastic order, which can be interpreted as a comparison of the outage probabilities, and do not exploit the full gamut of stochastic orders available [3]. In this chapter, we give a wide range of examples illustrating how different stochastic orders are appropriate for comparing systems using different metrics with analytical properties such as monotonicity, convexity, and

complete monotonicity, which shed light into the connections between performance metrics such as error rates and ergodic capacity. Additionally, we find the conditions for the preservation of inequalities satisfied by the averages of performance metrics of individual systems, when multiple such systems are combined. These may be combinations in parallel, in series, or otherwise, as may be seen in relay networks. Such a study permits the comparison of performance of systems, even in settings where closed-form expressions are not tractable.

In this chapter, $P_{\mathrm{e},i}(\rho)$ denotes the SER conditioned on the i^{th} symbol of the constellation being transmitted. $P_{\mathrm{e}}(\rho)$ denotes the SER averaged over the constellation. $\overline{P}_{\mathrm{e}}(\rho)$ denotes the average SER over a fading channel (where the averaging is over the constellation points and the fading channel statistics).

In the rest of the chapter, we illustrate the power of the stochastic ordering framework in comparing wireless channels and systems. We will investigate the convexity and complete monotonicity properties related to the error rate and capacity expressions in Section 3.2, which will facilitate comparing the average performances of systems by using (2.20) and (2.21). In Section 3.3 we identify commonly used channel distributions which are LT or convex ordered. Section 3.4 investigates the conditions under which these stochastic orders are preserved in complex systems where the performance of their constituent parts satisfy an order. Finally, relevant simulations to supplement the theoretical results are provided in Section 3.5.

3.2 Ordering of Average Error Rate and Ergodic Capacity Metrics

We now discuss how the tools provided by stochastic ordering theory can be used to compare the average error rate and ergodic capacity of two different systems.

3.2.1 Symbol Error Rate

It has been established in [12] that the error rate of binary signalling in the presence of noise with a uni-modal and differentiable PDF is a convex function of the SNR when maximum likelihood decoding is performed. Also, it is known that the instantaneous error probabilities of all one-dimensional and two-dimensional constellations with ML decoding in the presence additive white Gaussian noise (AWGN) is a convex function of the SNR [14]. In this section, we go one step further and establish the complete monotonicity of some two-dimensional modulation schemes, which will be useful in establishing inequalities between averaged performance metrics. It is well known that $\mathcal{Q}\left(\sqrt{x}\right)$ is c.m. [30], from which the complete monotonicity of the instantaneous error rate of the form $P_{\mathrm{e}}\ (\rho x) = a\mathcal{Q}\left(\sqrt{b\rho x}\right)$ easily follows, for $a, b > 0$. Here a and b are modulation dependent constants which can be chosen to get exact performance in some cases (e.g. $a = 1$, $b = 2$ for BPSK), or approximations in others ($a = 3/4, b = 4/5$ for 16-QAM). For the exact case, it follows from (2.20) that $\mathbb{E}\left[P_{\mathrm{e}}\ (\rho Y)\right] \leq \mathbb{E}\left[P_{\mathrm{e}}\ (\rho X)\right]$, for $\rho > 0$ whenever $X \leq_{\mathrm{Lt}}\ Y$.

We now establish, for the first time, the complete monotonicity of exact symbol error rates of square M-QAM and M-PSK modulations which are not in the form $P_{\mathrm{e}}\ (\rho x) = a\mathcal{Q}\left(\sqrt{b\rho x}\right)$, and offer sharper results than those mentioned above, since they do not rely on approximations.

The M-PSK symbol error rate is given by the following [2, pp.195]:

$$P_{\mathrm{e}}^{\ \mathrm{PSK}}\ (\rho x) = \frac{1}{\pi} \int_{0}^{(M-1)\pi/M} \exp\left(-\rho x \frac{\alpha_{\ \mathrm{PSK}}}{\sin^2\theta}\right) \mathrm{d}\theta\ , \tag{3.1}$$

where $\alpha_{\ \mathrm{PSK}} := \sin^2(\pi/M)$. After a change of variables, (3.1) can be expressed as the

Laplace transform of a positive function:

$$P_{\mathrm{e}}^{\ \mathrm{PSK}}\ (\rho x) = \frac{\sqrt{\alpha_{\ \mathrm{PSK}}}}{2\pi} \int_0^\infty e^{-\rho u x} \frac{I\left[u \geq \frac{\alpha_{\ \mathrm{PSK}}}{\sin^2\left((M-1)\frac{\pi}{M}\right)}\right]}{u\sqrt{u-\alpha_{\ \mathrm{PSK}}}} \mathtt{d}u \ , \tag{3.2}$$

which together with Bernstein's Theorem suggests that $P_{\mathrm{e}}^{\ \mathrm{PSK}}\ (\rho x)$ is c.m..

Consider now the square M-QAM error rate function [2, pp.195]:

$$P_{\mathrm{e}}^{\ \mathrm{QAM}}\ (\rho x) = a\mathcal{Q}\left(\sqrt{\alpha_{\ \mathrm{QAM}}\,\rho x}\right) - b\mathcal{Q}^2\left(\sqrt{\alpha_{\ \mathrm{QAM}}\,\rho x}\right) \ , \tag{3.3}$$

where $\alpha_{\ \mathrm{QAM}} := 3/(M-1)$, $a := 4(\sqrt{M}-1)/\sqrt{M}$ and $b := a^2/4$. Note that $0 \leq b \leq a$. We claim that (3.3) is c.m. for any a, b such that $b \leq a$. To see this, recall

$$\mathcal{Q}^k\left(\sqrt{x}\right) = \frac{1}{\pi} \int_0^{\pi/2k} \exp\left(-\frac{x}{2\sin^2\theta}\right) \mathtt{d}\theta \ , \tag{3.4}$$

for $k = 1, 2$ [2]. After a change of variables similar to (3.1), we obtain

$$P_{\mathrm{e}}^{\ \mathrm{QAM}}\ (\rho x) = \frac{\sqrt{\alpha_{\ \mathrm{QAM}}}}{2\pi} \int_0^\infty e^{-u\rho x} \left[\frac{aI\left[0.5 \leq u \leq 1\right]}{u\sqrt{2u-1}} + \frac{(a-b)I\left[u \geq 1\right]}{u\sqrt{2u-1}}\right] \mathtt{d}u \ , \tag{3.5}$$

which is also c.m. by Bernstein's theorem, since $b \leq a$. In conclusion, whenever $X \leq_{\mathrm{Lt}} Y$, $\mathbb{E}\left[P_{\mathrm{e}}\ (\rho Y)\right] \leq \mathbb{E}\left[P_{\mathrm{e}}\ (\rho X)\right]$ for all average SNR ρ, where $P_{\mathrm{e}}\ (\cdot)$ could be given by either (3.1) or (3.3). This follows from the definition of the LT order and the c.m. properties of instantaneous error rates of M-QAM or M-PSK modulations.

3.2.2 Ergodic Capacity

We now show that stochastic ordering of instantaneous channel power distributions implies that their ergodic channel capacities satisfy a corresponding inequality at all average SNRs. We begin with the case where only the receiver has channel status information (CSI).

Ergodic Capacity with Receive CSI only

The instantaneous capacity, conditioned on the channel power $X = x$ when only the receiver has CSI is given by $C(\rho x) = \log(1 + \rho x)$, where ρ is the average SNR. Since $dC(\rho x)/dx = \rho/(1 + \rho x)$ is c.m. in x, from (2.21) $X \leq_{\text{Lt}} Y$ implies that the ergodic capacities satisfy $\mathbb{E}_X [C(\rho X)] \leq \mathbb{E}_Y [C(\rho Y)]$ for $\rho \geq 0$. Recall that by (2.19), LT ordering of the channels X and Y can be interpreted as a comparison of the average error rates, when the instantaneous error rate is a decaying exponential. As a result, one can loosely say that *if the average error rates of two channels X and Y satisfy the inequality* (2.19) *at all SNRs, then so do the ergodic capacities.* Surprisingly, however, the converse is not true, as we now illustrate. Consider a Pareto-type distribution, which is appropriate for modeling the instantaneous SINR in the presence of interference [31] :

$$F_X(z) = \frac{z^\beta}{(1 + z^\beta)}, z > 0, \ \beta > 0 . \tag{3.6}$$

Using integration by parts and simplifying, we obtain

$$\mathbb{E}_X [C(\rho X)] = \int_0^\infty \frac{\rho}{(1 + \rho z)(1 + z^\beta)} \mathtt{d}z . \tag{3.7}$$

Taking the derivative with respect to β, it is seen that $\mathbb{E}_X [C(\rho X)]$ is a decreasing function of β, for $\rho > 0$. This shows that for $\beta^X \leq \beta^Y$, $\mathbb{E}_X [C(\rho X)] \geq \mathbb{E}_Y [C(\rho Y)]$ for $\rho > 0$. On the other hand, since $F_X(z) = z^\beta + o(z^\beta)$ near $z = 0$, the average symbol error rate for an exponential instantaneous error rate function satisfies $\mathbb{E}[\exp(-\rho X)] = (G_c \, \rho)^{-\beta} + o\left(\rho^{-\beta}\right)$, where G_c is the array gain and β is the diversity order [32]. Hence, as β increases, the high-SNR average error rate decreases, while the capacity also decreases at all SNR ρ! Interpreting the ergodic capacity as what is achievable by coding over an i.i.d. time-extension of the channel, we reach the conclusion that even though Y offers more diversity than X for an uncoded system,

the i.i.d. extension of X lends itself to more diversity than that of Y. To put it more simply, at high SNR, it is possible for one channel to be superior to another in terms of error rates in the absence of coding, while being inferior when the capacity achieving code is employed over both channels.

Channel Inversion and Delay-Limited Capacity

When CSI is available at the transmitter, it can be used for power adaptation. A simple, suboptimal approach is to "invert" the channel at the transmitter, so that effectively the receiver sees a non-fading AWGN channel. Such an approach is viable only when $\mathbb{E}\left[X^{-1}\right] < \infty$, leading to a finite average transmit power. This is the case whenever the channel offers a diversity order strictly greater than one. Channel inversion has the advantage that a channel code designed for the AWGN channel can be used effectively, and that the code length need not depend on the channel coherence time to average out the fading. This "delay-limited" approach [33] gives rise to an achievable rate given by

$$\overline{C}_{\mathrm{CI}}^{X}(\rho) = \log\left(1 + \frac{\rho}{\mathbb{E}\left[X^{-1}\right]}\right) . \tag{3.8}$$

Clearly, since $g(x) = x^{-1}$ is a c.m. function of x, $\mathbb{E}\left[X^{-1}\right] \geq \mathbb{E}\left[Y^{-1}\right]$, whenever $X \leq_{\mathrm{Lt}} Y$. This implies $\overline{C}_{\mathrm{CI}}^{X}(\rho) \leq \overline{C}_{\mathrm{CI}}^{Y}(\rho)$ for all ρ, since $\overline{C}_{\mathrm{CI}}^{X}(\rho)$ in (3.8) is a decreasing function of $\mathbb{E}\left[X^{-1}\right]$.

Optimal Power and Rate Adaptation (OA)

It is well known that CI is not optimal, since when the channel gain becomes arbitrarily small, the transmitter uses extremely high power. To overcome this limitation, the *optimal power and rate adaptation* scheme is proposed in [33], where water-filling across time is performed subject to an average transmit power constraint. The ca-

pacity so obtained over a channel with instantaneous SNR X is given by [33] :

$$\overline{C}_{\mathrm{OA}}^{X}(\rho) = \int_{z_t(\rho)}^{\infty} \log\left(\frac{z_t(\rho)}{z}\right) \mathtt{d}\left[1 - F_X\left(z\right)\right] \ , \tag{3.9}$$

where $z_t(\rho)$ is the signaling threshold, which is implicitly governed by the power constraint as follows:

$$\int_{z_t(\rho)}^{\infty} \left(\frac{1}{z_t(\rho)} - \frac{1}{z}\right) \mathtt{d}F_X\left(z\right) = \rho \ . \tag{3.10}$$

It can be shown that $X \leq_{\mathrm{Lt}} Y$ does not guarantee $\overline{C}_{\mathrm{OA}}^{X}(\rho) \leq \overline{C}_{\mathrm{OA}}^{Y}(\rho)$ for all ρ. However, in what follows, we will show that under the stronger assumption $X \leq_{\mathrm{st}} Y$, $\overline{C}_{\mathrm{OA}}^{X}(\rho) \leq \overline{C}_{\mathrm{OA}}^{Y}(\rho)$ for all ρ.

Using integration by parts on (3.10), it is observed that for $X \leq_{\mathrm{st}} Y$, we have $z_t^X(\rho) \geq z_t^Y(\rho)$. Now, integrating (3.9) by parts, under the assumptions that $\lim_{z\to\infty}(1 - F_X\left(z\right))\log(z/z_t^X(\rho)) = 0$ and $\lim_{z\to\infty}(1 - F_Y\left(z\right))\log(z/z_t^Y(\rho)) = 0$, it is seen that $\overline{C}_{\mathrm{OA}}^{X}(\rho) \leq \overline{C}_{\mathrm{OA}}^{Y}(\rho)$ for $\rho \geq 0$, since $z_t^X(\rho) \geq z_t^Y(\rho)$. Therefore, $X \leq_{\mathrm{st}} Y \Rightarrow \overline{C}_{\mathrm{OA}}^{X}(\rho) \leq \overline{C}_{\mathrm{OA}}^{Y}(\rho)$, for $\rho > 0$.

3.3 Ordering of Parametric Fading Distributions

We now proceed to show that commonly used parametric fading distributions are completely monotonic in the line of sight parameter with respect to LT and convex orders.

3.3.1 Nakagami Fading

Consider Nakagami fading model, where the envelope $\sqrt{X}$ is Nakagami and the channel power X is Gamma distributed [34], with PDF given by

$$f_X\left(x\right) = \frac{m^m}{\Gamma(m)} x^{m-1} \exp(-mx) \ , x \geq 0 \ . \tag{3.11}$$

Since $\mathbb{E}\left[\exp(-\rho X)\right] = (1+\rho/m)^{-m}$ is a decreasing function of m for each ρ, it follows that if the m parameters of two channel distributions satisfy $m^X \leq m^Y$, then $X \leq_{\text{Lt}} Y$, where X and Y have normalized Gamma distributions with parameters m^X and m^Y respectively. This shows that for example, all the performance metrics in (3.1) or (3.3) that are c.m. have averages over fading distributions that satisfy the inequality $\mathbb{E}_Y\left[P_{\text{e}}\ (\rho Y)\right] \leq \mathbb{E}_X\left[P_{\text{e}}\ (\rho X)\right]$ for all values of average SNR ρ. A similar claim with a reversal in the inequality can be made for the ergodic capacity metric. Note that the PDFs of X and Y in (3.11) are defined to satisfy $\mathbb{E}\left[X\right] = \mathbb{E}\left[Y\right] = 1$, independent of the fading parameter m. Hence, the improvements in error rate or ergodic capacity at all values of ρ with increased m is *not* due to an improvement in average SNR. A stronger convex ordering result can also be established. Since $\mathbb{E}\left[X\right] = \mathbb{E}\left[Y\right]$, $m^X \leq m^Y \Rightarrow Y \leq_{\text{cx}} X$ can be shown by using (2.15). We can summarize the results herein by using the terminology that the normalized Gamma distribution is monotonically increasing in m with respect to the orders $\leq_{\text{Lt}}$ and $\leq_{\text{cx}}$.

3.3.2 Ricean Fading

As in the Nakagami case, the Rice distribution will also be shown to be monotonic in the LoS parameter K with respect to the orders $\leq_{\text{Lt}}$ and $\leq_{\text{cx}}$. The instantaneous channel power distribution is given by

$$f_X(x) = (1+K)\exp(-K)\exp\left[-(K+1)x\right] I_0\left(2\sqrt{K(K+1)x}\right) , \qquad (3.12)$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind of order zero. Clearly, $\mathbb{E}\left[X\right] = 1$ is independent of K. Taking the Laplace transform of (3.12), we have $\mathbb{E}\left[\exp(-X\rho)\right] = (1+K)/(1+K+\rho)\exp\left[K\rho/(1+K+\rho)\right]$, which decreases with K for each ρ. This implies that, similar to the Nakagami case, increasing K without increasing the average SNR improves the average symbol error rate, ergodic capacity,

or any average metric obtained from a c.m. or c.m.d function. Thus, if $K^X \leq K^Y$ are the Ricean parameters of two channels, then $X \leq_{\text{Lt}} Y$. Similar to the Nakagami case, equation (2.15) can be used to establish a stronger claim that $Y \leq_{\text{cx}} X$.

In this specific Ricean context, similar results for the ergodic capacity are found in [35] and the references therein, in a more general MIMO setting. However, in these results, either the channel power increases with an increase in the LoS component, or only an asymptotically large number of antennas is considered.

3.4 Communication Systems Involving Multiple RVs

In the following discussion, we will consider systems involving multiple independent random channel coefficients and compare their performance in two different sets of channels, where the channel powers associated with the first system are denoted by $\mathbf{X} := [X_1, \ldots, X_M]$ while those of the second channel by $\mathbf{Y} := [Y_1, \ldots, Y_M]$. Toward this goal, we use the following result [3, pp. 97], which shows that LT ordering is preserved by multivariate functions that are c.m.d.:

Theorem 3.4.1. *Let $X_1, \ldots, X_M$ be independent and $Y_1, \ldots, Y_M$ also be independent. If $X_m \leq_{Lt} Y_m$ for $m = 1, \ldots, M$, then $g(X_1, \ldots, X_M) \leq_{Lt} g(Y_1, \ldots, Y_M)$ for all functions $g : \mathbb{R}^m \to \mathbb{R}^+$ such that for $m = 1, \ldots, M$, $(\partial/\partial x_m)\, g(x_1, \ldots, x_M)$ is c.m. in x_m, when all other variables are fixed.*

We now investigate the systems for which the combined instantaneous SNR is given by a function $g(\mathbf{x}) := g(x_1, \ldots, x_M)$, which satisfies the conditions of Theorem 3.4.1. Unless otherwise mentioned, we will assume throughout that $X_m \leq_{\text{Lt}} Y_m$ for $m = 1, \ldots, M$.

3.4.1 Maximum Ratio Combining

Consider a SIMO diversity combining system with M receive antennas which have complete CSI. If maximum ratio combining (MRC) is performed, conditioned on the channel powers $X_m = x_m$ for $m = 1, \ldots, M$, the instantaneous SNR at the output of the combiner is proportional to

$$g_{\text{MRC}}(\mathbf{x}) = \sum_{m=1}^{M} x_m \ , \tag{3.13}$$

which satisfies the conditions of Theorem 3.4.1 as easily seen by taking derivatives. Thus, we infer that when MRC is performed, $g_{\text{MRC}}(\mathbf{X}) \leq_{\text{Lt}} g_{\text{MRC}}(\mathbf{Y})$. Consequently, whenever $P_{\text{e}}(\rho x)$ is c.m. and $C(\rho x)$ is c.m.d., the average error rates satisfy $\mathbb{E}_{\mathbf{Y}}\left[P_{\text{e}}\left(\rho g_{\text{MRC}}(\mathbf{Y})\right)\right] \leq \mathbb{E}_{\mathbf{X}}\left[P_{\text{e}}\left(\rho g_{\text{MRC}}(\mathbf{X})\right)\right]$ for all ρ, and the ergodic capacities satisfy $\mathbb{E}_{\mathbf{Y}}\left[C(\rho g_{\text{MRC}}(\mathbf{Y}))\right] \geq \mathbb{E}_{\mathbf{X}}\left[C(\rho g_{\text{MRC}}(\mathbf{X}))\right]$, for all ρ.

3.4.2 Equal Gain Combining

Next, assume that the SIMO diversity system adopts equal gain combining (EGC) at the receiver. In this case, conditioned on the instantaneous channel powers $X_m = x_m$ for $m = 1, \ldots, M$, the instantaneous SNR at the combiner is proportional to

$$g_{\text{EGC}}(\mathbf{x}) = \frac{1}{M}\left(\sum_{m=1}^{M} \sqrt{x_m}\right)^2 \ . \tag{3.14}$$

The derivative $(\partial/\partial x_i)\, g_{\text{EGC}}(x_1, \ldots, x_M) = M^{-1}\left(\sum_{m=1}^{M} \sqrt{x_m}\right)/\sqrt{x_i}$ is a c.m. function of x_i, for $i = 1, \ldots, M$. Therefore, equal gain combining of a better set of branches results in a better system overall, as also expressed more rigorously after (3.13) in the MRC example.

3.4.3 Selection Combining

In contrast to the previous two examples, this example shows that even though the individual branch instantaneous channel powers are LT ordered, the combined SNR at the output of the combiner need not be LT ordered. For selection combining (SC), conditioning on the instantaneous channel powers $X_m = x_m$ for $m = 1, \ldots, M$, we have

$$g_{\text{SC}}(\mathbf{x}) = \max_m x_m \ , \tag{3.15}$$

which is not differentiable, and hence is not c.m.. In fact, $X_m \leq_{\text{Lt}} Y_m,\ m = 1, \ldots, M$ does not imply $\max_m X_m \leq_{\text{Lt}} \max_m Y_m$. We provide a simple counterexample in Section 3.5. This shows that even though channels Y_m provide better average error rates at all ρ than X_m, for $m = 1, \ldots, M$ for a SISO system, the composite SC channel does not.

3.4.4 Multi-hop Amplify and Forward (AF)

Consider a multi-hop system with M links subject to AWGN, where X_m is the channel power gain over the m^{th} link. It is assumed that the m^{th} node has channel information of the $(m-1)^{th}$ hop, for $m = 2, \ldots, M$, and the amplification factor for each node is the same. Conditioned on the instantaneous channel powers $X_m = x_m$ for $m = 1, \ldots, M$, the SNR at the destination in this case is proportional to [36]:

$$g_{\text{MH-AF}}(\mathbf{x}) = \left[\prod_{m=1}^{M} \left(1 + \frac{1}{x_m} \right) - 1 \right]^{-1} . \tag{3.16}$$

Taking the partial derivatives of $g_{\text{MH-AF}}(\mathbf{x})$ with respect to each x_m for $m = 1, \ldots, M$, it is seen that $g_{\text{MH-AF}}(\mathbf{x})$ satisfies the conditions of Theorem 3.4.1. Thus, $g_{\text{MH-AF}}(\mathbf{X}) \leq_{\text{Lt}} g_{\text{MH-AF}}(\mathbf{Y})$. As a result, the average error rates for the multi-hop AF system satisfy $\mathbb{E}_{\mathbf{Y}}\left[P_{\text{e}}\left(\rho g_{\text{MH-AF}}(\mathbf{Y})\right)\right] \leq \mathbb{E}_{\mathbf{X}}\left[P_{\text{e}}\left(\rho g_{\text{MH-AF}}(\mathbf{X})\right)\right]$, for $\rho > 0$. Importantly, a closed-form expression for the average performance of this system is not tractable for most

practical channel distributions. Despite this, it is still possible to compare the error rates of two otherwise identical systems systems with two sets of LT ordered channel powers at all average SNRs.

3.4.5 Multi-hop Channels with Decode and Forward

Consider an M-hop channel, where each terminal decodes a received symbol into a binary alphabet and forwards it over to the next terminal. Let the instantaneous error rate over the i^{th} link be given by $P_{\mathrm{e}_i}(\rho x_i)$, $i = 1, \ldots, M$, where we assume $0 \leq P_{\mathrm{e}_i}(x) \leq 1/2$ is c.m.. For convenience, we define $\mathbf{X}_{1:m} := [X_1, \ldots, X_m]$ and let $P_{\mathrm{e}_{1:m}}(\rho \mathbf{X}_{1:m})$ be the combined instantaneous error rate of the first $1 \leq m \leq M$ hops. We have the following:

Theorem 3.4.2. *Let $X_1, \ldots, X_M$ be independent, and likewise for $Y_1, \ldots, Y_M$. Suppose $X_m \leq_{Lt} Y_m$ for $m = 1, \ldots, M$. Then $\mathbb{E}_{\mathbf{X}_{1:m}}\left[P_{e_{1:m}}(\rho \mathbf{X}_{1:m})\right] \geq \mathbb{E}_{\mathbf{Y}_{1:m}}\left[P_{e_{1:m}}(\rho \mathbf{Y}_{1:m})\right]$, $m = 1, \ldots, M$.*

Proof. For any m, viewing the m-hop channel as a series cascade of the first $m-1$ hops and the m^{th} hop, we have the following:

$$\begin{aligned} P_{\mathrm{e}_{1:m}}(\rho \mathbf{X}_{1:m}) = {} & P_{\mathrm{e}_{1:m-1}}(\rho \mathbf{X}_{1:m-1})\left(1 - P_{\mathrm{e}_m}(\rho x_m)\right) \\ & + \left(1 - P_{\mathrm{e}_{1:m-1}}(\rho \mathbf{X}_{1:m-1})\right) P_{\mathrm{e}_m}(\rho x_m) \ , \end{aligned} \tag{3.17}$$

for $m = 2, \ldots, M$. To prove the theorem, we will use induction. Clearly, Theorem 3.4.2 holds for $m = 1$. Taking expectation of both sides of (3.17), we have

$$\begin{aligned} \mathbb{E}_{\mathbf{X}_{1:m}}\left[P_{\mathrm{e}_{1:m}}(\rho \mathbf{X}_{1:m})\right] = {} & \mathbb{E}_{\mathbf{X}_{1:m-1}}\left[P_{\mathrm{e}_{1:m-1}}(\rho \mathbf{X}_{1:m-1})\right]\left(1 - \mathbb{E}_{X_m}\left[P_{\mathrm{e}_m}(\rho X_m)\right]\right) \\ & + \left(1 - \mathbb{E}_{\mathbf{X}_{1:m-1}}\left[P_{\mathrm{e}_{1:m-1}}(\rho \mathbf{X}_{1:m-1})\right]\right) \mathbb{E}_{X_m}\left[P_{\mathrm{e}_m}(\rho X_m)\right] \ . \end{aligned} \tag{3.18}$$

We have $\mathbb{E}_{\mathbf{X}_{1:m-1}}\left[P_{\mathrm{e}_{1:m-1}}(\rho\mathbf{X}_{1:m-1})\right] \geq \mathbb{E}_{\mathbf{Y}_{1:m-1}}\left[P_{\mathrm{e}_{1:m-1}}(\rho\mathbf{Y}_{1:m-1})\right]$ by the induction hypothesis, and $\mathbb{E}_{\mathbf{X}_{1:m}}\left[P_{\mathrm{e}_{1:m}}(\rho\mathbf{X}_{1:m})\right] \geq \mathbb{E}_{\mathbf{Y}_{1:m}}\left[P_{\mathrm{e}_{1:m}}(\rho\mathbf{Y}_{1:m})\right]$ follows because $P_{\mathrm{e}_m}(\cdot)$ is c.m. and $X_m \leq_{\mathrm{Lt}} Y_m$. The theorem then follows because the RHS of (3.18) is of the form $P_1(1-P_2)+P_2(1-P_1)$, which is an increasing function of both P_1 and P_2, since $0 \leq P_1 \leq 1/2$, $0 \leq P_2 \leq 1/2$. □

Note that Theorem 3.4.2 and its proof carry over when each hop adopts M-ary modulation as well, provided that ρ is large enough to ensure $0 \leq \mathbb{E}_{\mathbf{X}_{1:m}}\left[P_{\mathrm{e}_{1:m}}(\rho\mathbf{X}_{1:m})\right] \leq 1/2$.

3.4.6 Post Detection Combining

Consider an M-antenna post-detection combining (PDC) scheme, where the instantaneous symbol error rate on the m^{th} branch is $P_{\mathrm{e}_m}(\rho x_m)$ and is c.m. as in the previous example. The instantaneous probability of error of the PDC system is given by[1] [37] :

$$P_{\mathrm{e}_{1:M}}(\rho\mathbf{X}_{1:M}) = \sum_{k=\frac{M+1}{2}}^{M} \sum_{\mathcal{S}_k} \left(\prod_{i\in\mathcal{S}_k} P_{\mathrm{e}_i}(\rho x_i)\right)\left(\prod_{j\in\mathcal{S}_k^c}\left(1-P_{\mathrm{e}_j}(\rho x_j)\right)\right), \tag{3.19}$$

where $\mathcal{S}_k$ is a set running over all subsets of $\{1,\ldots,M\}$ with k elements. Taking expectation with respect to $\mathbf{X}_{1:M}$, which is assumed to have independent components, we have,

$$\begin{aligned}&\mathbb{E}_{\mathbf{X}_{1:m}}\left[P_{\mathrm{e}_{1:m}}(\rho\mathbf{X}_{1:m})\right]\\ &= \sum_{k=\frac{M+1}{2}}^{M} \sum_{\mathcal{S}_k} \left(\prod_{i\in\mathcal{S}_k} \mathbb{E}_{X_i}\left[P_{\mathrm{e}_i}(\rho X_i)\right]\right)\left(\prod_{j\in\mathcal{S}_k^c}\left(1-\mathbb{E}_{X_j}\left[P_{\mathrm{e}_j}(\rho X_j)\right]\right)\right).\end{aligned} \tag{3.20}$$

[1] We assume M is odd. Extensions to even M are straight-forward by adding a tie breaker term to (3.19).

Clearly, the average error rate is an increasing function of $\mathbb{E}_{X_m}\left[P_{\mathrm{e}_m}(\rho X_m)\right]$, since it is not possible to get improved performance by increasing the average error rate on any particular link. This shows that when $X_m \leq_{\mathrm{Lt}} Y_m$, for $m = 1, \ldots, M$, and $P_{\mathrm{e}_m}(\rho x)$ is c.m., so that the average error rates of PDC satisfy $\mathbb{E}_X\left[P_{\mathrm{e}_m}(\rho X_m)\right] \geq \mathbb{E}_Y\left[P_{\mathrm{e}_m}(\rho Y_m)\right]$ for $\rho > 0$, it follows that $\mathbb{E}_{\mathbf{X}}\left[P_{\mathrm{e}_{1:M}}(\rho \mathbf{X}_{1:M})\right] \geq \mathbb{E}_{\mathbf{Y}}\left[P_{\mathrm{e}_{1:M}}(\rho \mathbf{Y}_{1:M})\right]$ for $\rho > 0$.

3.4.7 Generalized Multi-branch Multi-hop AF Cooperative Relay Networks

We now consider the generalized relay structure illustrated Fig. 3.1, which consists of M independent branches, each involving N_m relays, for $m = 1, \ldots, M$, which assist the direct link between the source S and the destination D by performing amplify and forward (AF). It is assumed that all the links are impaired by AWGN with fixed variance. This model requires the branches to communicate through mutually orthogonal channels, so that M independent copies are available at the destination which performs MRC (using combining coefficients given in [38]). Although approximate expressions for the error rate have been obtained for the case of Ricean fading in [38], closed-form expressions are intractable.

Note that the two-hop fixed AF relay, which finds frequent application in cooperative diversity literature [38] and illustrated in Fig. 3.2 is a special case of this general relay, with $M = 1$ and $N_m = 1$. Thus, the forthcoming results obtained for the general case apply for the two-hop relay as well. We now show that the exact average symbol error rate can be compared over a number of fading distributions where the pairs of channel powers are LT ordered. To this end, we show that the output SNR of the MRC combiner at the destination satisfies the conditions of Theorem 3.4.1. Let $X_{0,0}$ denote the channel power on the direct link, and $X_{m,n}$ the channel power at the n^{th} hop on link m. Since the destination performs MRC, the instantaneous output SNR is the sum of individual end-to-end branch SNRs, each of which are given by

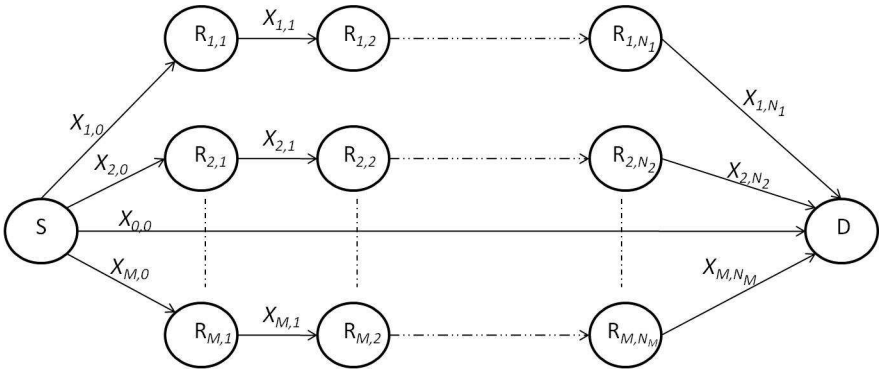

Figure 3.1: Multi-branch multi-hop cooperative relay network. $\mathrm{R}_{m,1} \ldots \mathrm{R}_{m,N_m}$ represent the relays on the m^{th} link from the source S to the destination D. The corresponding instantaneous channel power gains are denoted as $X_{m,0} \ldots X_{m,N_m}$.

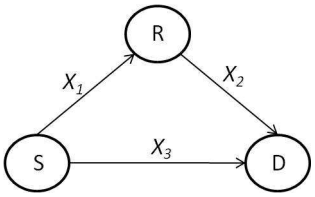

Figure 3.2: Two hop AF cooperative relay network.

(3.16). Thus, conditioned on $X_{m,n} = x_{m,n}$ for $m = 0, \ldots, M$ and $n = 0, \ldots, N_m$, and defining $g_{\text{MB-MH-AF}}(\mathbf{X}) := g_{\text{MB-MH-AF}}\left(x_{0,0}, x_{1,0}, \ldots, x_{1,N_1}, \ldots, x_{M,N_M}\right)$, we have

$$g_{\text{MB-MH-AF}}(\mathbf{X}) = \sum_{m=1}^{M} \left[\prod_{n=1}^{N_m} \left(1 + \frac{1}{x_{m,n}}\right) - 1 \right]^{-1} + x_{0,0} \,. \tag{3.21}$$

As shown in the arguments following (3.16), the summand in the RHS of (3.21) has a c.m. derivative in each variable. Combining this with Theorem 3.4.1, we have $\mathbb{E}_{\mathbf{Y}}\left[P_{\text{e}}\left(\rho g_{\text{MB-MH-AF}}(\mathbf{Y})\right)\right] \leq \mathbb{E}_{\mathbf{X}}\left[P_{\text{e}}\left(\rho g_{\text{MB-MH-AF}}(\mathbf{X})\right)\right]$ for $\rho > 0$.

3.4.8 Combined Multipath Fading and Shadowing

It is well known that the effect of shadow fading on the instantaneous SNR distribution can be modeled as a product of a shadowing random variable with a multipath fading random variable [2]. Let $X_1 \leq_{\text{Lt}} Y_1$ be the two multipath fading SNR distributions, and $X_2 \leq_{\text{Lt}} Y_2$ be the two shadowing distributions. Then, from Theorem 3.4.1, it follows that the composite RV satisfies $X_1 X_2 \leq_{\text{Lt}} Y_1 Y_2$,

since $g(x_1, x_2) = x_1 x_2$ has a c.m. derivative in each variable. We conclude that $\mathbb{E}_{X_1,X_2}\left[P_{\mathrm{e}}\ (\rho X_1 X_2)\right] \geq \mathbb{E}_{Y_1,Y_2}\left[P_{\mathrm{e}}\ (\rho Y_1 Y_2)\right],\ \forall \rho$, whenever $P_{\mathrm{e}}\ (\cdot)$ is c.m.. Such conclusions can be drawn even in cases where the composite distribution of $X_1 X_2$ or $Y_1 Y_2$ cannot be written in closed-form.

3.4.9 Systems with non-Gaussian Channel Noise

In this discussion, we assume the following system model:

$$Z = \sqrt{\rho X} S + W\ , \tag{3.22}$$

where for simplicity, $S \in \{-1, 1\}$, ρX is the instantaneous SNR, ρ the average SNR, and W is non-Gaussian noise.

Gaussian Mixture

In this model, W represents compound Gaussian noise (also called Gaussian mixture), which can be written as $W = \sqrt{A}G$, where A is a positive valued RV, which represents the scale of G, and $G \sim \mathcal{N}(0, 1)$. Such a formulation is possible for symmetric alpha-stable noise, Middleton class-A noise, as well as other compound Gaussian noise distributions. The error rate conditioned on the channel power $X = x$ is given by

$$P_{\mathrm{e}}\ (\rho x) = \mathbb{E}_A\left[\mathcal{Q}\left(\sqrt{\frac{2\rho x}{A}}\right)\right]\ , \tag{3.23}$$

which is a c.m. function of x as can be verified by differentiating inside the expectation with respect to x. Using (2.20), this shows that when $X \leq_{\mathrm{Lt}} Y$ then the average error rates satisfy $\mathbb{E}_X\left[P_{\mathrm{e}}\ (\rho X)\right] \geq \mathbb{E}_Y\left[P_{\mathrm{e}}\ (\rho Y)\right]$, even for mixed (compound) Gaussian noise. Similar results can also be shown to hold for noise distributions such as the Laplace distribution which cannot be expressed as a compound Gaussian.

Bounded Noise

Recall the system model from (3.22). If $|W| \leq C$ for some constant C, almost surely then $F_W(x) = 1$ for $x \geq C$ and $1 - F_W\left(\sqrt{2x}\right) = 0$ for $x^2/2 \geq C$. It is clear from Bernstein's theorem that a function, such as $1 - F_W\left(\sqrt{2x}\right)$ with bounded support cannot be c.m.. From this, we can conclude that if the noise is bounded, it is possible for two SNR distributions to be LT ordered, although $\mathbb{E}_Y\left[P_{\mathrm{e}}\left(\rho Y\right)\right]$ need not be less than $\mathbb{E}_X\left[P_{\mathrm{e}}\left(\rho X\right)\right]$ for all $\rho > 0$. This negative result emphasizes the effect of the noise distribution in claims of ordering and concludes our discussion of systems with non-Gaussian noise.

3.5 Simulations

We now corroborate our theoretical results using Monte-Carlo simulations. For ease of notation, we define $\overline{P}_{\mathrm{e}}^{\,\mathrm{X}}(\rho) := \mathbb{E}_X\left[P_{\mathrm{e}}\left(\rho X\right)\right]$ to denote the average error rates of SISO systems operated in the channel power X. Also, we use $\overline{P}_{\mathrm{e}}^{\,\mathbf{X}}(\rho) := \mathbb{E}_{\mathbf{X}}\left[P_{\mathrm{e}}\left(\rho g\left(\mathbf{X}\right)\right)\right]$ to represent the average error rates of systems involving multiple channel power coefficients.

One of the central results of Section 3.2.2 is that it is possible for one channel to be superior to another (in terms of error rates) at high SNR in the absence of coding, while being inferior when the capacity achieving code is used over both channels. This is illustrated in Fig. 3.3, which shows the comparative error rate performance of DPSK employed over an interference dominated fading channel with Pareto type distributed instantaneous SINR (having parameters $\beta^X = 2$ and $\beta^Y = 5$. Clearly, since $\overline{P}_{\mathrm{e}}^{\,\mathrm{X}}(\rho) < \overline{P}_{\mathrm{e}}^{\,\mathrm{Y}}(\rho)$ for $\rho < -0.5$ dB and vice-versa for $\rho > -0.5$ dB, the system with channel power X is not better than that with channel power Y at every average SNR. On the other hand, Fig. 3.4, shows that the ergodic capacity of the system

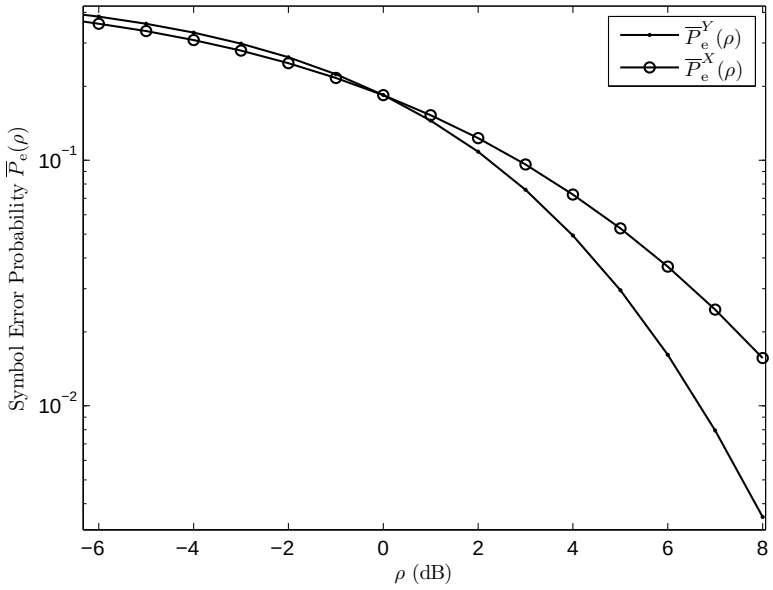

Figure 3.3: Error probability comparison of DPSK modulation, under two different fading scenarios with LT ordered Pareto-type SIR distributions, using parameters $\beta^X = 2$ and $\beta^Y = 5$.

with instantaneous channel X is consistently larger than that when operated in the channel Y with parameter $\beta^Y = 5$. Figures 3.5, 3.6 and 3.7 show the performance of diversity combining schemes such as MRC, EGC and SC with $L = 3$ branches over two sets of i.i.d Ricean fading channels with parameters $K^X = 2$ and $K^Y = 5$. Note that from Subsection 3.3.2, $X_m \leq_{\text{Lt}} Y_m$ for $m = 1, 2, 3$. The trend observed in the performance analysis curves obtained herein can be equivalently obtained for any other sets of LT ordered channel power random variables, using any modulation scheme whose error rate is a c.m. function of the channel power.

In Fig. 3.5, we demonstrate that LT ordering of the instantaneous SNR distributions for the individual branches can be used to compare average error rates when MRC is performed at the receiver. For $L = 3$ receive diversity branches, it is observed that the error rate of BPSK in the channel with instantaneous SNR ρY is consistently less than that in the channel with instantaneous SNR ρX, which agrees with the fact

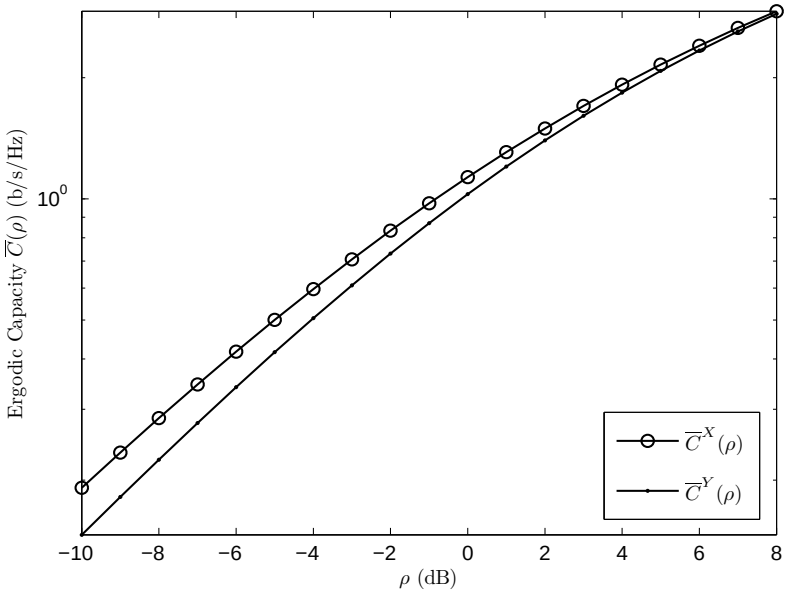

Figure 3.4: Ergodic capacity comparison of two different fading scenarios with LT ordered Pareto-type SNR distributions, using parameters $\beta^X = 2$ and $\beta^Y = 5$. $\overline{C}^X(\rho)$ ($\overline{C}^Y(\rho)$) represents the ergodic capacity in the channel power X (Y).

that since the channel power for Ricean fading is c.m. in K, $X_m \leq_{\text{Lt}} Y_m$, for $m = 1, 2, 3$, and hence $\overline{P}_{\text{e}}^{\mathbf{Y}}(\rho) \leq \overline{P}_{\text{e}}^{\mathbf{X}}(\rho)$ for $\rho > 0$.

Figure 3.6 illustrates that when $X_m \leq_{\text{Lt}} Y_m$, for $m = 1, 2, 3$, we get $\overline{P}_{\text{e}}^{\mathbf{Y}}(\rho) \leq \overline{P}_{\text{e}}^{\mathbf{X}}(\rho)$ for $\rho > 0$ for the case of EGC employing BPSK. The error rate curves help demonstrate that fading channels with larger Ricean parameters offer smaller error rates than those with smaller Ricean parameters at all values of average SNR ρ when EGC is used, as predicted in Subsection 3.4.2. Such a conclusion is not present in the literature due to the unavailability of a closed-form expression for the average error rate of coherent EGC in Ricean channels, which is applicable in all SNR regimes [2].

The comparative performance of SC using DPSK symbols is shown in Fig. 3.7. It is evident that although the individual branch SNRs are LT ordered, $\overline{P}_{\text{e}}^{\mathbf{Y}}(\rho) \geq \overline{P}_{\text{e}}^{\mathbf{X}}(\rho)$, for $\rho < -0.4$ dB, while $\overline{P}_{\text{e}}^{\mathbf{Y}}(\rho) \leq \overline{P}_{\text{e}}^{\mathbf{X}}(\rho)$, for $\rho \geq -0.4$ dB. This cross-over point is clearly depicted in Fig. 3.7 using a linear scale for the error rate axis, since

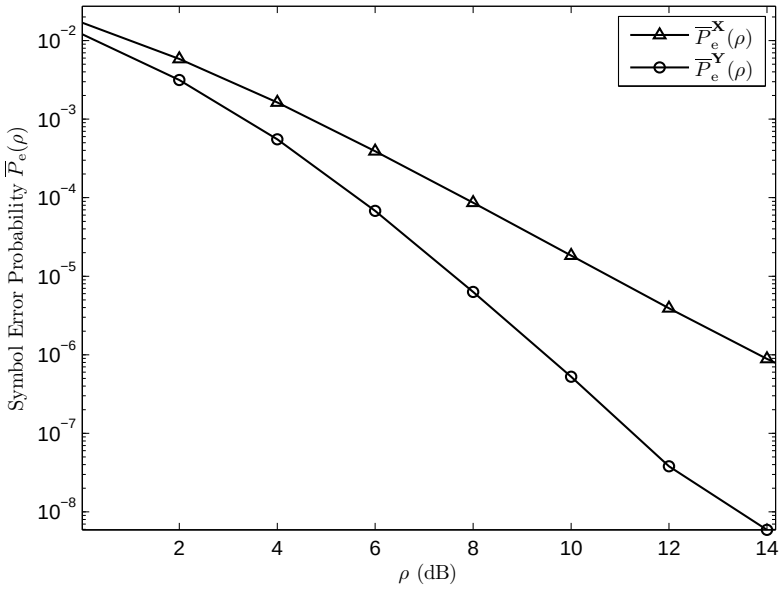

Figure 3.5: Error rate comparison of maximum ratio combining using $L = 3$ antennas with BPSK. $\overline{P}_{\mathrm{e}}^{\,\mathbf{X}}(\rho)$ corresponds to the average symbol error rate under Ricean fading with parameter $K^X = 2$ and $\overline{P}_{\mathrm{e}}^{\,\mathbf{Y}}(\rho)$ to the average symbol error rate under Ricean fading with parameter $K^Y = 5$.

it is more easily discernible compared to the conventional log scale. Hence, selection combining of a better set of channels (in terms of error rates) need not yield a better system overall, at low SNR.

The performance of a multi-hop amplify and forward relay is studied in Fig. 3.8. We assume the model described in Section 3.4.4 with $M = 3$ relays under two different Ricean fading scenarios, one with parameter $K^X = 2$ and the other with $K^Y = 5$. It is observed that the average symbol error rate of $\mathbf{Y}$ is consistently less than that of $\mathbf{X}$ at all SNRs. This, due to the fact that $X_m \leq_{\mathrm{Lt}} Y_m,\ m = 1, 2, 3 \Rightarrow \overline{P}_{\mathrm{e}}^{\,\mathbf{Y}}(\rho) \leq \overline{P}_{\mathrm{e}}^{\,\mathbf{X}}(\rho),\ \forall \rho$.

Fig. 3.9 illustrates the comparative performance of an uncoded BPSK system over an additive compound Gaussian noise channel subject to two different Ricean fading effects modeled using parameters $K^X = 2$ and $K^Y = 5$. We show that $\overline{P}_{\mathrm{e}}^{\,\mathrm{Y}}(\rho) \leq$

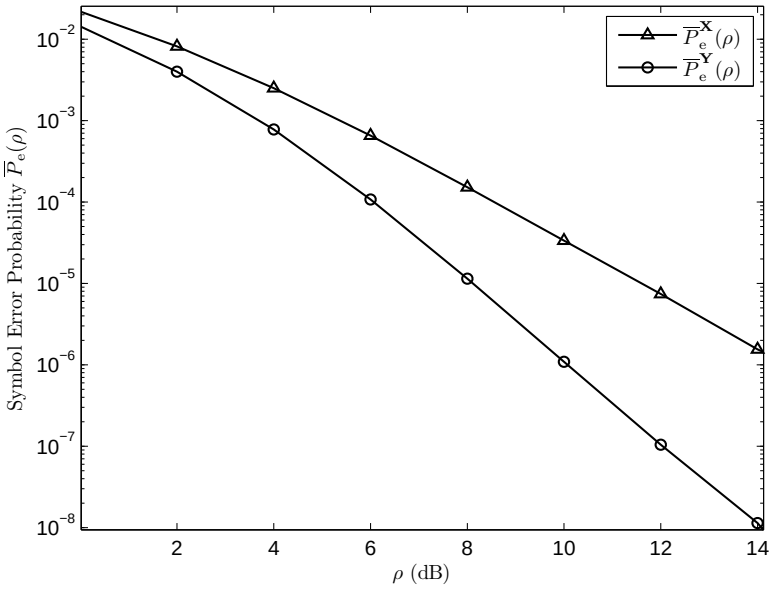

Figure 3.6: Error rate comparison of equal gain combining using $L = 3$ antennas with BPSK. $\overline{P}_{\mathrm{e}}^{\,\mathbf{X}}(\rho)$ corresponds to the average error rate under Ricean fading with parameter $K^X = 2$ and $\overline{P}_{\mathrm{e}}^{\,\mathbf{Y}}(\rho)$ to the average symbol error rate under Ricean fading with parameter $K^Y = 5$.

$\overline{P}_{\mathrm{e}}^{\,\mathrm{X}}(\rho)$ for all $\rho > 0$, when the noise follows a symmetric alpha-stable distribution with a characteristic exponent of 1.6. This shows that LT ordering results apply to systems with compound Gaussian noise, since an alpha-stable RV can be written as $\sqrt{A}G$, where $G \sim \mathcal{CN}(0, 1)$ and A is a positively skewed alpha-stable RV [37]. Such results are not found in literature, since a closed-form expression for the average error rate of BPSK under Ricean fading with symmetric alpha-stable noise is analytically intractable. In fact, even for the special case of $K = 1$ i.e. Rayleigh fading, a closed-form expression valid in the asymptotic high SNR regime is known [37].

In direct contrast to the results for the compound-Gaussian noise case, LT ordering of channel powers does not imply that the average error rate performance for noise with bounded support will satisfy the corresponding inequality at all SNR. In fact, as depicted in Fig. 3.10, where the unit-variance noise is assumed to be uniformly

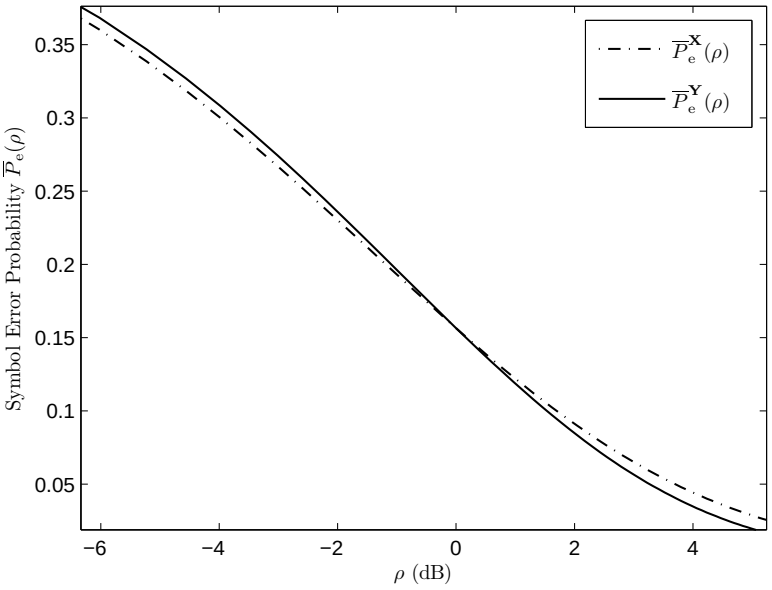

Figure 3.7: Error rate comparison of selection combining using $L = 3$ antennas with DPSK. $\overline{P}_{\rm e}^{\,\mathbf{X}}(\rho)$ corresponds to the average symbol error rate under Ricean fading with parameter $K^X = 2$ and $\overline{P}_{\rm e}^{\,\mathbf{Y}}(\rho)$ to the average symbol error rate under Ricean fading with parameter $K^Y = 5$.

distributed on $[-\sqrt{3}, \sqrt{3}]$, it is observed that for $\rho < 2.6$ dB, $\overline{P}_{\rm e}^{\,\rm X}(\rho) \leq \overline{P}_{\rm e}^{\,\rm Y}(\rho)$, while the opposite holds for $\rho > 2.6$ dB. This corroborates the claim of Subsection 3.4.9, which states that LT ordering of channel powers does not imply that the average error rates satisfy $\overline{P}_{\rm e}^{\,\rm Y}(\rho) \leq \overline{P}_{\rm e}^{\,\rm X}(\rho)$ for all $\rho > 0$, under noise with finite support.

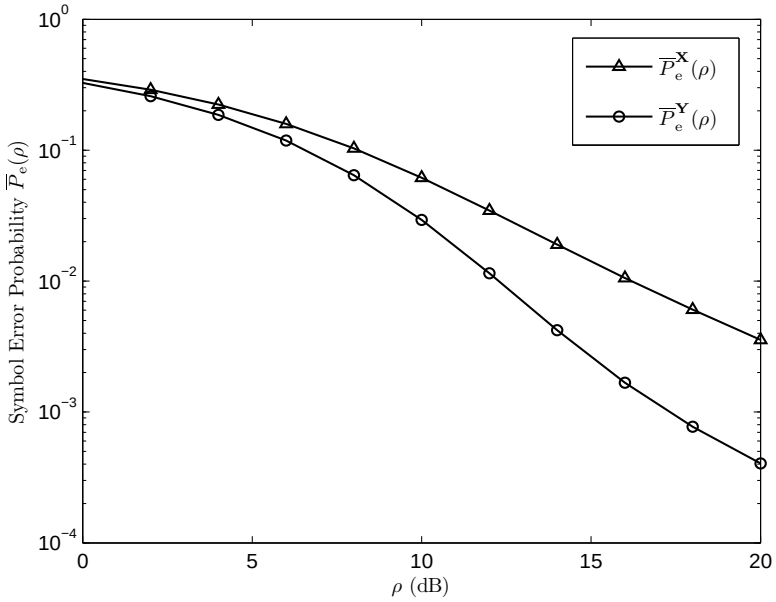

Figure 3.8: Error rate comparison of $M = 3$ hop amplify-forward relay with BPSK under Ricean fading. $\overline{P}_{\mathrm{e}}^{\,\mathbf{X}}(\rho)$ corresponds to the average symbol error rate under Ricean fading with parameter $K^X = 2$ and $\overline{P}_{\mathrm{e}}^{\,\mathbf{Y}}(\rho)$ to the average error rate under Ricean fading with parameter $K^Y = 5$.

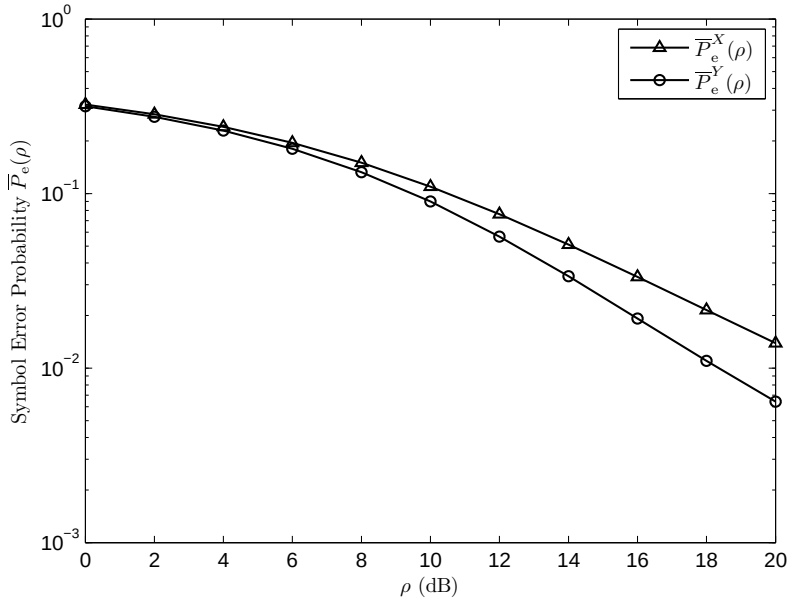

Figure 3.9: Performance comparison of BPSK in compound Gaussian noise (normalized symmetric alpha-stable distribution with characteristic exponent 1.6). $\overline{P}_{\mathrm{e}}^{\,\mathrm{X}}(\rho)$ corresponds to the average symbol error rate under Ricean fading with parameter $K^X = 2$ and $\overline{P}_{\mathrm{e}}^{\,\mathrm{Y}}(\rho)$ corresponds to the average symbol error rate under Ricean fading with parameter $K^Y = 5$.

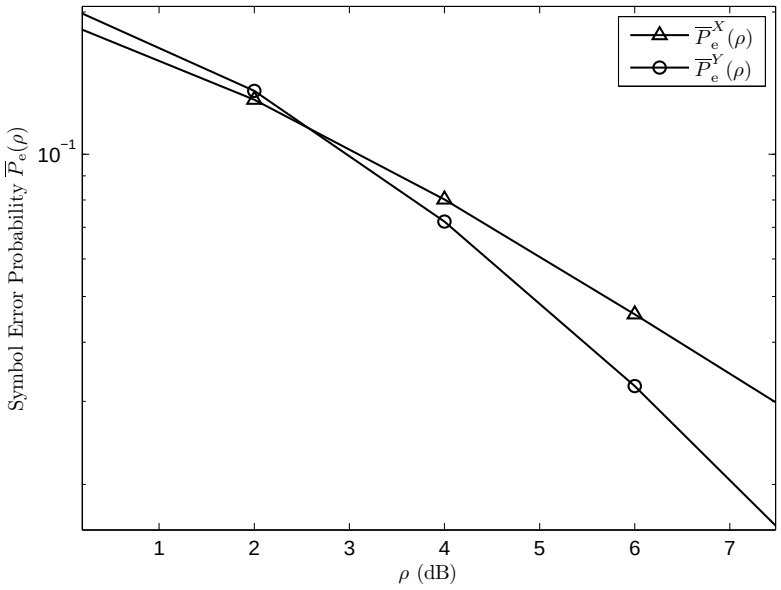

Figure 3.10: Performance comparison of BPSK in noise with finite support (symmetric uniformly distributed noise with unit variance). $\overline{P}_{\mathrm{e}}^{\mathrm{X}}(\rho)$ corresponds to the average symbol error rate under Ricean fading with parameter $K^X = 2$ and $\overline{P}_{\mathrm{e}}^{\mathrm{Y}}(\rho)$ corresponds to the average error rate under Ricean fading with parameter $K^Y = 5$.

Chapter 4

SYMBOL ERROR RATES OF ARBITRARY CONSTELLATIONS IN AWGN - REPRESENTATION AND PROPERTIES

4.1 Motivation and Literature Survey

The complete monotonicity of a performance metric such as the symbol error rate plays a pivotal role in ordering two fading channels with respect to this metric, as described in Chapter 3. Consequently, a study of the complete monotonicity properties of the SER of arbitrary constellations is useful, although it is largely absent in the literature. Specifically, investigations into the properties of the SER of one-dimensional and two-dimensional constellations have revealed the convexity of the SER with respect to the signal-to-noise ratio (SNR) under impairing additive white Gaussian noise (AWGN) [12, 14]. Some special cases of two dimensional constellations such as M-ary phase shift keying (M-PSK) and M-ary quadrature amplitude modulation (M-QAM) have SERs which are known to be completely monotone functions of the SNR [39, 30], which is a stronger condition than convexity. On the other hand, constellations of dimensionality greater than two (which we refer to as "higher dimensional constellations" henceforth) have found practical applications in satellite communications [40, 41] and more recently, in optical communications [42, 43]. Investigations of the convexity properties of the SER of such constellations are relatively scarce in the literature. It is known that the second derivative of the SER of a constellation of dimensionality greater than two is non-negative at sufficiently high SNR [14]. Although this result is a general one, it does not provide a conclusive means of determining whether a given arbitrary constellation has a convex SER or not. For

certain higher dimensional constellations, analytical expressions for the SER have been derived in the literature, which can be used to deduce the convexity of the SER. For example, the class of constellations of dimensionality $2, 3, 4$ and 5 described in [44] can be verified to have convex SERs under AWGN, by differentiation of the analytical SER expressions given in [44]. On the other hand, verifying if a SER is completely monotone is difficult, even if a closed form expression for the SER is available. It has been established recently, that if the rank of a constellation matrix is at most two, then it will have a completely monotone SER under AWGN with ML detection [45]. However, there has been no investigation into the complete monotonicity properties of the SER for higher dimensional constellations, which can be used to simplify the expressions for average SERs over fading channels, and to establish useful comparisons of average SERs of a system under two different fading channels, as described in Chapter 3 and [39].

In this work, it is shown that the SER of an arbitrary multi-dimensional constellation in the absence of coding, when impaired by additive independent and identically distributed (i.i.d.) Gaussian noise under maximum likelihood (ML) detection can be represented as a product of a power of the SNR and a completely monotone function of SNR. This result also generalizes to SERs under compound Gaussian noise, which includes many non-Gaussian noise distributions such as Middleton class-A noise [15] and symmetric alpha-stable noise [16]. The SER of an arbitrary multi-dimensional constellation is shown to be completely monotone if the constellation matrix has a rank of one or two. Since complete monotonicity implies convexity, the SER is a convex function of the SNR, provided that the constellation matrix has a rank of one or two. For constellations matrices whose rank is greater than two, it is shown that the complete monotonicity of the SER depends on the constellation geometry and choice of prior probabilities. This work also describes a novel stochastic order for

fading distributions, which can be used to order the average SERs of arbitrary multi-dimensional complex constellations over quasi-static fading channels, and generalizes the existing Laplace transform order on random variables.

The rest of the chapter is organized as follows: Section 2.1 surveys the relevant mathematical background. Section 4.2 describes the system model. The result on the representation of the SER of an arbitrary multi-dimensional constellation is detailed in Section 4.3. In Section 4.4, the applications such as (i) ordering the average SERs of constellations with c.m. SERs over two different fading channels, and (ii) comparing the average SER of an arbitrary constellation using a newly proposed stochastic order, are discussed.

4.2 System Model

In this chapter, the transmission of an N-dimensional uncoded baseband signal $\mathbf{s}$ through AWGN is considered, which is described as follows:

$$\mathbf{y} = \mathbf{s} + \mathbf{z} \ , \tag{4.1}$$

where the transmitted symbol $\mathbf{s} \in \mathbb{R}^N$ is chosen from $\mathcal{S} := \{\mathbf{s}_1, \ldots, \mathbf{s}_M\}$. The *constellation matrix* corresponding to $\mathcal{S}$ is defined as $\mathbf{S} := [\mathbf{s}_1, \ldots, \mathbf{s}_M]$, and the *reduced dimension* of $\mathbf{S}$ is defined as the rank of $\mathbf{S}$, which is denoted by N^*. In (4.1), the noise is assumed to be $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, (1/\rho)\mathbf{I})$, $\rho > 0$. When the signal energy is normalized as $M^{-1}\sum_{i=1}^{M} ||\mathbf{s}_i||^2 = 1$, then $M^{-1}\sum_{i=1}^{M} ||\mathbf{s}_i||^2 / \mathbb{E}\left[\mathbf{z}^{\mathrm{T}}\mathbf{z}\right] = \rho$ represents the average SNR per dimension. At the receiver, the ML detector under AWGN is assumed, where the detected symbol $\hat{\mathbf{s}}$ is given by:

$$\hat{\mathbf{s}} = \underset{\mathbf{s}\in\mathcal{S}}{\operatorname{argmin}} \, ||\mathbf{y} - \mathbf{s}||^2 \ , \tag{4.2}$$

which is the ML detector for white Gaussian noise. Assuming that the origin of the coordinate system is shifted to $\mathbf{s}_i$, the Voronoi region of $\mathbf{s}_i$, denoted by $\mathcal{K}_i$ is given by

$$\mathcal{K}_i := \{\mathbf{x} \in \mathbb{R}^N | \mathbf{A}_i \mathbf{x} \leq \mathbf{b}_i\}, \tag{4.3}$$

where $\mathbf{A}_i \in \mathbb{R}^{F \times N}$, where $F \leq M$, and the j^{th} row of $\mathbf{A}_i$ is $\mathbf{a}_{j,i}^{\mathrm{T}} = (\mathbf{s}_j - \mathbf{s}_i)^{\mathrm{T}} / ||\mathbf{s}_j - \mathbf{s}_i||$, while the j^{th} element of $\mathbf{b}_i$ is $b_{j,i} = ||\mathbf{s}_j - \mathbf{s}_i|| / 2$. It is assumed that (4.3) is a non-redundant description of $\mathcal{K}_i$. The minimum distance $d_{\min}$ of the constellation is defined through its square as $d_{\min}^2 := \min_{\mathbf{s}_i, \mathbf{s}_j \in \mathcal{S}; \mathbf{s}_j \neq \mathbf{s}_i} ||\mathbf{s}_i - \mathbf{s}_j||^2 = 4 \min_{i,j} b_{j,i}^2$ [46].

The probability of error $P_{\mathrm{e},i}(\rho)$, given that $\mathbf{s}_i$ is transmitted is given by

$$P_{\mathrm{e},i}(\rho) := \left(\frac{\rho}{2\pi}\right)^{N/2} \int_{\mathbb{R}^N - \mathcal{K}_i} \exp\left[-\frac{\rho}{2} \mathbf{x}^{\mathrm{T}} \mathbf{x}\right] \mathtt{d}\mathbf{x} \,, \tag{4.4}$$

where $\mathcal{S}_1 - \mathcal{S}_2 := \{\mathbf{x} \in \mathcal{S}_1 | \mathbf{x} \notin \mathcal{S}_2\}$ is the relative complement of $\mathcal{S}_2$ in $\mathcal{S}_1$. The probability of error averaged across all possible transmitted symbols is given by

$$P_{\mathrm{e}}(\rho) = \sum_{i=1}^{M} \Pr[\mathbf{s} = \mathbf{s}_i] P_{\mathrm{e},i}(\rho) \;, \tag{4.5}$$

where $\Pr[\mathbf{s} = \mathbf{s}_i]$ represents the *a priori* probability of transmitting $\mathbf{s}_i$.

For any constellation $\mathbf{S}$, there exists an equivalent full rank constellation $\mathbf{S}^*$, which has the same SER as that of $\mathbf{S}$. Such a definition is useful in developing a representation for the SER of a multidimensional constellation.

Definition 4.2.1 (Reduced Constellation)**.** *Let $\mathbf{S} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ be a singular value decomposition of $\mathbf{S}$, where $\mathbf{U} \in \mathbb{R}^{N \times N}$, $\mathbf{\Sigma} \in \mathbb{R}^{N \times M}$ and $\mathbf{V} \in \mathbb{R}^{M \times M}$, and the diagonal matrix consisting of the first N^* rows and first N^* columns of $\mathbf{\Sigma}$ contains the non-zero singular values of $\mathbf{S}$. Then the $N^* \times M$ matrix $\mathbf{S}^*$ given by the first N^* rows of $\mathbf{\Sigma}\mathbf{V}^T$ is defined as the reduced constellation of $\mathbf{S}$.*

By definition, the number of rows of $\mathbf{S}^*$ is no greater than that of $\mathbf{S}$. In addition, $\mathbf{S}^*$ can be shown to have the same SER as $\mathbf{S}$. To see this, recall that the SER of

the minimum distance detector depends only on the distance between the column vectors of $\mathbf{S}$ and the Frobenious norm of $\mathbf{S}$. It then suffices to show that the columns of $\mathbf{S}^*$ have the same distance properties and norm as that of $\mathbf{S}$. To this end, observe that $\mathbf{\Sigma V}^{\mathrm{T}} = \mathbf{U}^{\mathrm{T}}\mathbf{S}$ has the same distance and norm properties as that of $\mathbf{S}$, since it is an orthogonal transformation on $\mathbf{S}$. Further, by construction, the last $N - N^*$ rows of $\mathbf{\Sigma V}^{\mathrm{T}}$ are zeros. Hence, the distance and norm properties of $\mathbf{\Sigma V}^{\mathrm{T}}$ and $\mathbf{S}^*$ are identical. In addition, since $\mathbf{z}$ is AWGN, multiplying $\mathbf{z}$ by an orthogonal matrix does not change its statistics. It thus follows that the SER of $\mathbf{S}^*$ and $\mathbf{S}$ are identical. To conclude the discussion of reduced constellations, an example of a constellation and its reduced constellation is provided. Consider $\mathcal{S} = \{\mathbf{s}_1, \mathbf{s}_2\}$, where $\mathbf{s}_1 = [\sqrt{0.5}\ \sqrt{0.5}]^{\mathrm{T}}$, and $\mathbf{s}_2 = [-\sqrt{0.5}\ \ -\sqrt{0.5}]^{\mathrm{T}}$. Using the definition of the reduced constellation, it is straightforward to see that the reduced constellation corresponding to $\mathcal{S}$ is the BPSK constellation set, and therefore the SER of $\mathcal{S}$ is identical to that of BPSK.

4.3 Symbol Error Rates of Multi-Dimensional Constellations

Throughout the chapter, the focus is to obtain a functional characterization of the SER of a multidimensional constellation, rather than to obtain bounds or closed-form expressions for the SER. Such a characterization can be used to uncover its convexity and complete monotonicity properties.

4.3.1 Symbol Error Rates Under AWGN

To begin with, it is assumed the the transmitted symbol is a real vector, and the additive noise is white Gaussian. For constellations with reduced dimension $N^* = 1$, the SER of the detector (4.2) under AWGN can be seen to be a positive linear combination of c.m. functions of the form $\mathcal{Q}\left(\sqrt{2\rho\eta}\right), \eta > 0$, which is c.m.. The functional structure of the SER of constellations with $N^* \geq 2$ is addressed in Theorem

4.3.1, whose proof requires a result from the combinatorial geometry literature, which is stated next.

Lemma 4.3.1. *Let $\mathcal{P}$ be an N-dimensional polyhedron in $\mathbb{R}^N$ with F facets. If $\mathcal{P}$ contains the origin in its interior, then $\mathbb{R}^N$ admits the decomposition $\mathcal{X} := \{\mathcal{D}_{f,q_f}\}_{f,q_f}$, $f = 1, \ldots, F$, $q_f = 1, \ldots, Q_f$, where $\mathcal{D}_{f,q_f}$ are N-dimensional simplicial cones, and f can be viewed as an index of the facets of $\mathcal{P}$.*

Proof. The proof of this lemma rests on the fact that it is possible to decompose $\mathbb{R}^N$ into a set $\mathcal{F}$ consisting of polyhedral cones $\mathcal{C}_1, \ldots, \mathcal{C}_F$, using the facets of an N-dimensional polyhedron $\mathcal{P} \in \mathbb{R}^N$, which contains the origin in its interior [24, pp. 192]. Since every polyhedral cone admits a decomposition into N-dimensional simplicial cones [25, Lemma 1.40], it is possible to decompose each $\mathcal{C}_f$ into a set of N-dimensional simplicial cones $\{\mathcal{D}_{f,q_f}\}_{q_f}$, for $f = 1, \ldots, F$. Consequently, $\mathbb{R}^N$ admits a decomposition into N-dimensional simplicial cones, given by $\{\mathcal{D}_{f,q_f}\}_{f,q_f}$. □

In other words, using the facets of an N-dimensional polyhedron $\mathcal{P} \subseteq \mathbb{R}^N$ which contains the origin, it is possible to decompose $\mathbb{R}^N$ into a collection of N-dimensional simplicial cones. In what follows, the representation theorem is stated.

Theorem 4.3.1. *For a constellation $\mathcal{S} \subseteq \mathbb{R}^N$, whose reduced constellation is $\mathcal{S}^*$ and reduced dimension is $N^* \geq 2$, the SER of the detector (4.2) under AWGN admits the representation*

$$P_e(\rho) = \rho^p f_{cm}(\rho) \,, \tag{4.6}$$

where $f_{cm}(\rho)$ is c.m., and $p \geq N^*/2 - 1$. In (4.6), the representing function of $f_{cm}(\rho)$ satisfies $\mu(u) = 0$ when $u < d_{min}^2/4$, and $\mu(u) \geq 0$ otherwise, where d_{min} is the minimum distance of the constellation.

Proof. Throughout this appendix, we work with the reduced constellation $\mathbf{S}^*$, with rank N^*. Recall the AWGN system model (4.1). To obtain an expression for the symbol error rate averaged over all constellation points $P_{\mathrm{e}}(\rho)$, we first evaluate $P_{\mathrm{e},i}\,(\rho)$ given by (4.4), and then use (4.5). For the sake of simplicity, we assume that the Voronoi region of $\mathbf{s}_i^* \in \mathbf{S}^*$ is a polytope. The following proof can easily be extended to cases when $\mathcal{K}_i$ is an unbounded polyhedron, by assuming an additional facet $\mathbf{c}_0\mathbf{x} \leq 1$, which turns $\mathcal{K}_i$ into a polytope [24, pp. 75], and subsequently taking the limit of $P_{\mathrm{e}}(\rho)$ so obtained as $\mathbf{c}_0 \to 0$.

We begin with an outline of the proof. In general, evaluating (4.4) is not straightforward, since $\mathcal{K}_i$ is a polytope. However, the Gaussian integral in (4.4) can be simplified, if the region of integration is of the form $\mathcal{Z} - \tilde{\mathcal{Z}}$, where $\mathcal{Z}$ is an N-dimensional simplicial cone, and $\tilde{\mathcal{Z}}$ is the intersection of a halfspace and $\mathcal{Z}$. To this end, we show that the Voronoi region of $\mathbf{s}_i^*$ has a dimension of N^*, so that Lemma 4.3.1 can be used, and thus $P_{\mathrm{e},i}\,(\rho)$ in (4.4) can be rewritten as a sum of integrals over regions of the form $\mathcal{Z} - \tilde{\mathcal{Z}}$, where $\mathcal{Z}$ and $\tilde{\mathcal{Z}}$ are as defined above. Each of these integrals when expressed in hyperspherical coordinates yields a canonical structure, which can be algebraically manipulated to obtain (4.6) in the Theorem. In order to show that $N^*/2-1$ in (4.6) is the smallest exponent for which the Theorem holds, an argument involving complete monotonicity of order α is provided towards the end of this appendix.

In what follows, we present the details of the proof. Let $\mathcal{K}_i$ be a non-redundant description of the Voronoi region of $\mathbf{s}_i^* \in \mathcal{S}^*$, for $i = 1, \ldots, M$. With a slight abuse of notation, we are dropping the superscript $*$ from $\mathcal{K}_i$, to simplify the notation. First, we show that $\mathcal{K}_i$ satisfies the conditions of Lemma 4.3.1. To this end, for any set of affinely independent points in space (such as $\mathcal{S}^*$), the dimension of the Voronoi region corresponding to each point is equal to the dimension of the affine hull of the set (N^*, when the set of points is $\mathcal{S}^*$) [47, p. 232]. Therefore, $\mathcal{K}_i$ is an N^*-dimensional polytope

in $\mathbb{R}^{N^*}$. Also, since the origin of the coordinate system is shifted to $\mathbf{s}_i^*$, $\mathbf{0} \in \mathcal{K}_i$. Thus, $\mathcal{K}_i$ satisfies the conditions of Lemma 4.3.1. As a result, using Lemma 4.3.1 we obtain a set $\mathcal{X}_i := \{\mathcal{D}_{f,q,i}\}_{q,f}$, which is a decomposition of $\mathbb{R}^{N^*}$ into N^*-dimensional simplicial cones (see Fig. 1). Clearly, every $\mathbf{x} \notin \mathcal{K}_i$ satisfies $\mathbf{x} \in \mathcal{D}_{f,q,i} - \mathcal{K}_i$ for some $\{q, f\}$. Let the number of facets of $\mathcal{K}_i$ be F_i. It now follows that

$$P_{\mathrm{e},i}(\rho) = \sum_{f=1}^{F_i} \sum_{q=1}^{Q_{f,i}} J_{f,q,i}(\rho) \,, \tag{4.7}$$

where

$$J_{f,q,i}(\rho) = \left(\frac{\rho}{2\pi}\right)^{N^*/2} \int\limits_{\mathcal{D}_{f,q,i} - \mathcal{K}_i} \exp\left(-\frac{\rho}{2} \sum_{k=1}^{N^*} x_k^2\right) \mathtt{d}x_1 \ldots \mathtt{d}x_{N^*} \,. \tag{4.8}$$

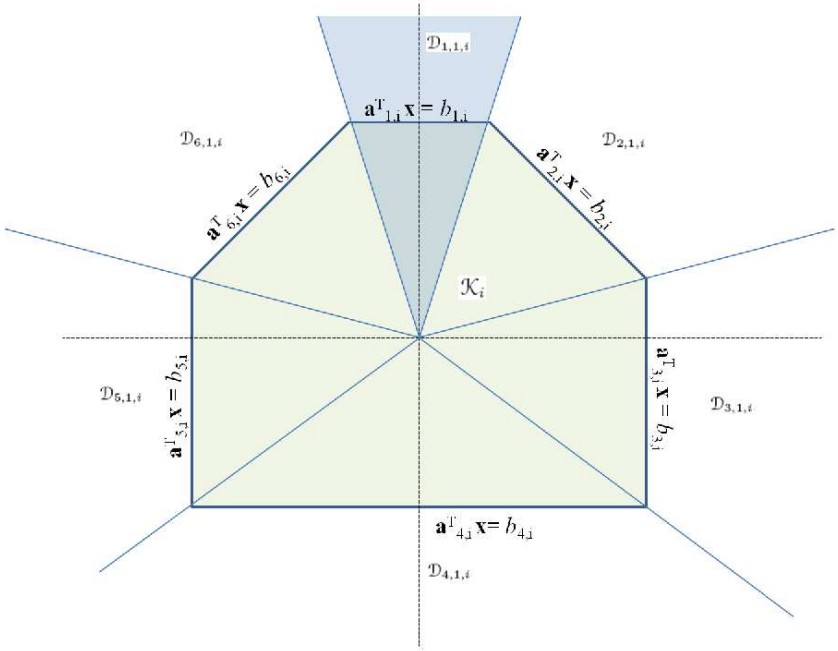

Figure 4.1: Voronoi region for signal point $\mathbf{s}_i^* \in \mathcal{S}^*$, where the reduced dimension of the constellation $\mathcal{S}$ is $N^* = 2$. The origin of the coordinate axes is shifted to $\mathbf{s}_i^*$. Using Lemma 1, $\mathbb{R}^{N^*}$ is decomposed into a collection of 2-dimensional simplicial cones $\mathcal{X}_i := \{\mathcal{D}_{1,1,i}, \ldots, \mathcal{D}_{6,1,i}\}$ using the facets of $\mathcal{K}_i$.

In order to simplify the integral in (4.8), we switch to the hyperspherical coordinate system [48], which is a generalization of the spherical coordinate system to higher

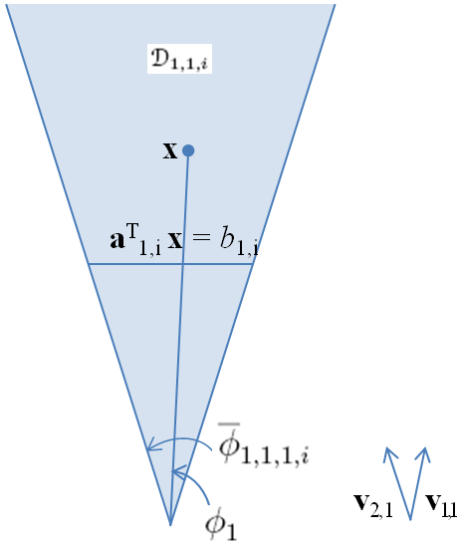

Figure 4.2: The 2-dimensional simplicial cone $\mathcal{D}_{1,1,i}$ obtained using Lemma 1 with $\mathcal{K}_i$, represented in hyperspherical coordinates.

dimensions. In this system, $\mathbf{x} \in \mathbb{R}^{N^*}$ is uniquely represented as $[r, \phi_1, \ldots, \phi_{N^*-1}]$, where $r = ||\mathbf{x}||$, and ϕ_k is the angle between $\mathbf{x}$ and the k^{th} edge of $\mathcal{D}_{f,q,i}$, $k = 1, \ldots, N^* - 1$. More precisely, let $\mathbf{v}_{k,q}$ define the unit vector in the direction of the k^{th} edge of $\mathcal{D}_{f,q,i}$, for $k = 1, \ldots, N^*$. Then, $\phi_k = \cos^{-1}(\mathbf{x}^{\mathrm{T}}\mathbf{v}_{k,q}/\,||\mathbf{x}||)$, $k = 1, \ldots, N^* - 1$ (See Fig. 4.2 for a two-dimensional example.).

Next, we obtain the region of integration in (4.8) in hyperspherical coordinates. For any $\mathbf{x} \in \mathcal{D}_{f,q,i} - \mathcal{K}_i$ represented by $[r, \phi_1, \ldots, \phi_{N^*-1}]$, the parameter r must satisfy

$$\overline{r}_{f,q,i}(\boldsymbol{\phi}) \leq r \leq \infty \,, \tag{4.9}$$

where $\overline{r}_{f,q,i}(\boldsymbol{\phi})$ is the distance of the point $\mathbf{x} \in \{\mathbf{x} | \mathbf{a}^{\mathrm{T}}_{f,i}\mathbf{x} = b_{f,i}\}$ from the origin. An expression for $\overline{r}_{f,q,i}(\boldsymbol{\phi})$ can be found by representing the hyperplane $\mathbf{a}^{\mathrm{T}}_{f,i}\mathbf{x} = b_{f,i}$ in hyperspherical coordinates, using the inverse hyperspherical transform relations

$$\begin{aligned} x_k &= r \cos\phi_k \prod_{k_1=1}^{k-1} \sin\phi_{k_1}, k = 1, \ldots, N^* - 1 \,, \\ x_{N^*} &= r \prod_{k_1=1}^{N^*-1} \sin\phi_{k_1} \,, \end{aligned} \tag{4.10}$$

and solving for r as a function of $\boldsymbol{\phi} := [\phi_1, \ldots, \phi_{N^*-1}]$. Thus, we get

$$\overline{r}_{f,q,i}(\boldsymbol{\phi}) = \frac{b_{f,i}}{\sum_{k=1}^{N^*-1} a_{k,f,i} \cos\phi_k \prod_{k_1=1}^{k-1} \sin\phi_{k_1} + a_{N^*-1,f,i} \prod_{k_1=1}^{N^*-1} \sin\phi_{k_1}} , \tag{4.11}$$

In (4.11), $a_{k,f,i}$ is the k^{th} element of $\mathbf{a}_{f,i}^{\mathrm{T}}$. Also, for any $\mathbf{x} \in \mathcal{D}_{f,q,i} - \mathcal{K}_i$, it is seen that ϕ_k must be at least 0 radians (if $\mathbf{x} = \alpha \mathbf{v}_{_{k,q}}, \alpha > 0$), and at most $\overline{\phi}_{k,f,q,i}$, which is the angle between $\mathbf{v}_{_{N^*,q}}$ and $\mathbf{v}_{_{k,q}}, k = 1, \ldots, N^* - 1$. In other words,

$$0 \leq \phi_{_k} \leq \cos^{-1}\left(\mathbf{v}_{_{N^*,q}}^{\mathrm{T}} \mathbf{v}_{_{k,q}}\right) =: \overline{\phi}_{k,f,q,i} . \tag{4.12}$$

It is useful to note that $\overline{\phi}_{k,f,q,i} \leq \pi$, since it is the angle between any two edges of the simplicial cone $\mathcal{D}_{f,q,i}$, which is at most π.

Thus, (4.8) can now be reformulated in hyperspherical coordinates, with the limits of integration given by (4.9) and (4.12) as

$$J_{f,q,i}(\rho) = \left(\frac{\rho}{2\pi}\right)^{N^*/2} \int_0^{\overline{\phi}_{N^*-1,f,q,i}} \cdots \int_0^{\overline{\phi}_{1,f,q,i}} \int_{\overline{r}_{f,q,i}(\boldsymbol{\phi})}^{\infty} r^{N^*-1} s(\boldsymbol{\phi}) e^{-\rho r^2/2} \mathtt{d}r \mathtt{d}\boldsymbol{\phi} , \tag{4.13}$$

where $s(\boldsymbol{\phi}) := \prod_{k=1}^{N^*-2} \sin^{N^*-k-1}\phi_k$ arises from the Jacobian of the transformation, and $\mathtt{d}\boldsymbol{\phi} = \mathtt{d}\phi_{_1} \ldots \mathtt{d}\phi_{_{N^*-1}}$. Substituting $u = r^2/2$ in (4.13), and changing the order of integration, we get

$$P_{\mathrm{e},i}(\rho) = \rho^{N^*/2} \int_0^{\infty} e^{-\rho u} \tilde{\mu}_i(u) \mathtt{d}u , \tag{4.14}$$

where $\tilde{\mu}_i(u)$ is given by

$$\tilde{\mu}_i(u) := \frac{1}{2\pi^{N^*/2}} \sum_{f=1}^{F_i} \sum_{q=1}^{Q_{f,i}} \sum_{l=1}^{L} \int_{\underline{\theta}_{N^*-1,f,q,i,l}}^{\overline{\theta}_{N^*-1,f,q,i,l}} \cdots \int_{\underline{\theta}_{1,f,q,i,l}}^{\overline{\theta}_{1,f,q,i,l}} s(\boldsymbol{\phi}) u^{N^*/2-1} \cdot I\left[\overline{r}_{f,q,i}(\boldsymbol{\phi})^2 \leq u\right] \mathtt{d}\boldsymbol{\phi} . \tag{4.15}$$

In (4.15), L is the number of convex intervals of $[\phi_1, \ldots, \phi_{N^*-1}]$ obtained after changing the order of integration, since the inverse function of $\overline{r}_{f,q,i}(\boldsymbol{\phi})^2$ is not unique. We now show that $\tilde{\mu}_i(u) \geq 0$, which, together with Bernstein's Theorem implies that the integral in (4.14) is equivalent to a c.m. function of ρ. To this end, observe that the integrand in (4.15) is non-negative for $u \geq 0$, because $s(\boldsymbol{\phi}) \geq 0$ for $\overline{\phi}_{k,f,q,i} \in [0, \pi], k = 1, \ldots, N^* - 1$. Consequently, the result obtained after the $N^* - 1$ fold integration in (4.15) is also non-negative. Thus, $\tilde{\mu}_i(u)$, which is a scaled version of a sum of non-negative integrals, is also non-negative. Therefore, through Bernstein's Theorem, we can assert that $P_{\mathrm{e},i}(\rho) = \rho^{N^*/2}\tilde{f}_{\mathrm{cm},i}(\rho)$, where $\tilde{f}_{\mathrm{cm},i}(\rho)$ is a c.m. function.

Now, using (4.14) in (4.5), we get

$$P_{\mathrm{e}}(\rho) = \rho^{N^*/2} \int_0^\infty e^{-\rho u}\tilde{\mu}(u)\mathtt{d}u \ , \tag{4.16}$$

where $\tilde{\mu}(u) := \sum_{i=1}^{M} \Pr[\mathbf{s} = \mathbf{s}_i]\tilde{\mu}_i(u)$. Thus, $P_{\mathrm{e}}(\rho) = \rho^{N^*/2}\tilde{f}_{\mathrm{cm}}(\rho)$, where $\tilde{f}_{\mathrm{cm}}(\rho)$ is c.m. through Bernstein's Theorem, because $\tilde{\mu}(u) \geq 0$ as it is a positive linear combination of non-negative functions $\mu_i(u), i = 1, \ldots, M$.

Next, we strengthen the representation (4.16) by showing that $\rho^{\alpha}\tilde{f}_{\mathrm{cm}}(\rho)$ is c.m. for $\alpha = 1$. To this end, recall from Section 2.1.1 that the necessary and sufficient condition for a c.m. function to be c.m. of order 1 is that its representing function be nonnegative and increasing. This is indeed the case for $\tilde{\mu}(u)$. Thus, we have just showed that $\tilde{f}_{\mathrm{cm}}(\rho)$ is c.m. of order $\alpha = 1$. Denoting $f_{\mathrm{cm}}(\rho) := \rho\tilde{f}_{\mathrm{cm}}(\rho)$, which we have just showed to be c.m., we get a stronger representation for the SER using (4.16) as follows

$$P_{\mathrm{e}}(\rho) = \rho^{\frac{N^*}{2}-1} f_{\mathrm{cm}}(\rho) \ , \tag{4.17}$$

where $f_{\mathrm{cm}}(\rho)$ is c.m..

Next, the support of the representing function of $f_{\mathrm{cm}}(\rho)$ is investigated. Let $\mu(u)$

be the representing function of $f_{\mathrm{cm}}(\rho)$. Integration by parts on (4.16) yields

$$\mu(u) = \sum_{i=1}^{M}\sum_{f=1}^{F_i}\sum_{q=1}^{Q_{f,i}}\sum_{l=1}^{L}\int_{\underline{\theta}_{N^*-1,f,q,i,l}}^{\overline{\theta}_{N^*-1,f,q,i,l}} \cdots \int_{\underline{\theta}_{1,f,q,i,l}}^{\overline{\theta}_{1,f,q,i,l}} s(\boldsymbol{\phi})\frac{\Pr[\mathbf{s}=\mathbf{s}_i]}{2\pi^{N^*/2}}$$
$$\left((N^*/2-1)u^{\frac{N^*}{2}-2}I\left[\overline{r}_{f,q,i}(\boldsymbol{\phi})^2 \le u\right] + u^{\frac{N^*}{2}-1}I\left[\overline{r}_{f,q,i}(\boldsymbol{\phi})^2 = u\right]\right)\mathtt{d}\boldsymbol{\phi}\ , \quad (4.18)$$

which is zero if $u < \min\limits_{q,f,i}\inf\limits_{\boldsymbol{\phi}}\overline{r}_{f,q,i}(\boldsymbol{\phi})^2$, and non-negative otherwise. Recalling the expression for $\overline{r}_{f,q,i}(\boldsymbol{\phi})$ from (4.11), it is immediately seen that $\min\limits_{q,f,i}\inf\limits_{\boldsymbol{\phi}}\overline{r}_{f,q,i}(\boldsymbol{\phi})^2 = \min_{f,i} b_{f,i}^2$. Observing that $\min_{f,i} b_{f,i}^2 = d_{\min}^2/4$, we conclude that the support of μ is contained in $[d_{\min}^2/4, \infty)$.

This concludes the proof of the Theorem. □

To prove Theorem 4.3.1, we work with a reduced constellation $\mathbf{S}^*$ that is full rank. This is needed since Lemma 4.3.1 is used in Theorem 4.3.1, where $\mathcal{P}$ is a Voronoi region with dimension N^*. This highlights the need to work with the reduced constellation and dimension. The proof of this Theorem can be viewed as a generalization of the method adopted in [49] to obtain an expression for the SER of arbitrary two-dimensional constellations under AWGN.

As an example to corroborate Theorem 4.3.1, consider the square M-QAM constellation under AWGN. For this constellation, it is known that $d_{\min} = \sqrt{2}$ and $P_{\mathrm{e}}(\rho) = \omega_1\mathcal{Q}\left(\sqrt{\eta\rho}\right) - \omega_2\mathcal{Q}^2\left(\sqrt{\eta\rho}\right)$, with $\omega_1 = 4(\sqrt{M}-1)/\sqrt{M}, \eta = 3/(M-1)$ and $\omega_2 = \omega_1^2/4$ [2]. In this case, $P_{\mathrm{e}}(\rho)$ can be represented in the form (4.6), where $p = 0$ (since $N^* = 2$), and the representing function of $f_{\mathrm{cm}}(\rho)$ is given by

$$\mu(u) = \frac{\sqrt{\eta}}{2\pi}\left[\frac{\omega_1 I[0.5 \le u \le 1]}{u\sqrt{2u-1}} + \frac{(\omega_1-\omega_2)I[u \ge 1]}{u\sqrt{2u-1}}\right]\ , \quad (4.19)$$

which is zero when $u < 0.5$ and non-negative when $u \ge 0.5$, because $\omega_1 - \omega_2 > 0$. Further, since $p = 0$, it is obvious that the square M-QAM constellation is an example of a constellation which has a c.m. SER.

According to Theorem 4.3.1, the SER of every constellation can be written as $P_{\mathrm{e}}(\rho) = \rho^p f_{\mathrm{cm}}(\rho)$, where $p = N^*/2 - 1$ and f_{cm} is c.m.. However, it is possible that the SER of some constellations admit a representation of the form (4.6), where the exponent of ρ is less than $N^*/2 - 1$. As a result, constellations for which this exponent is zero have c.m. SERs. The following corollary of Theorem 4.3.1 establishes a necessary and sufficient condition for the SER of a constellation to be c.m..

Corollary 4.3.1. *The SER of $\mathcal{S}$ using the detector (4.2) under AWGN is c.m. if its reduced dimension satisfies $N^* \leq 2$. Conversely, let $\mu(u) = \int_0^\infty e^{\rho u} \rho^{-(N^*/2-1)} P_e(\rho) \mathtt{d}\rho$, $\hat{\mu}(u) = \int_0^u \mu(u-v) v^{-1/2} \mathtt{d}v$ and $r = \lceil N^*/2 - 2 \rceil$. If $N^* > 2$ and N^* is even (odd), the SER is c.m. if and only if $\mu(u)$ ($\hat{\mu}(u)$) is r times differentiable, and $\mathtt{d}^k \mu(u)/\mathtt{d}u^k \geq 0$ ($\mathtt{d}^k \hat{\mu}(u)/\mathtt{d}u^k \geq 0$) for $0 \leq k \leq r$, and $\mathtt{d}^r \mu(u)/\mathtt{d}u^r$ ($\mathtt{d}^r \hat{\mu}(u)/\mathtt{d}u^r$) is increasing and continuous.*

Proof. The direct part of the corollary for constellations with $N^* = 1$ is immediate, since $P_{\mathrm{e}}(\rho)$ is a positive linear combination of functions of the form $\mathcal{Q}\left(\sqrt{2\rho\eta}\right), \eta > 0$, which is known to be c.m. (see e.g. [39]). Also, the complete monotonicity of the SER for the case of $N^* = 2$ is straightforward from Theorem 4.3.1. To see the converse, we use the necessary and sufficient condition for a function $f(\rho)$ to be c.m. of order α, as described in Section 2.1.1. For the case when N^* is even, we assume $f(\rho) = f_{\mathrm{cm}}(\rho)$, where f_{cm} is as defined in Theorem 4.3.1, and $\alpha = N^*/2 - 1$. For the case of odd N^*, we assume $f(\rho) = \rho^{-1/2} f_{\mathrm{cm}}(\rho)$, and $\alpha = \lceil N^*/2 - 1 \rceil$. The converse thus easily follows. □

In other words, Corollary 4.3.1 states that the complete monotonicity of the SER for constellations with $N^* \leq 2$ does not depend on the geometry of the constellation. However, for constellations with higher reduced dimensions, the complete monotonicity of the SER depends on the differentiability of the representing function

corresponding to $f_{\mathrm{cm}}(\rho)$, which is a function of the constellation geometry and the a-priori probabilities. Although Corollary 4.3.1 is applicable to any constellation, it is not easy to obtain the equivalent set of conditions on the constellation geometry and prior probabilities under which the SER of a constellation with $N^* > 2$ is c.m., and this is posed as an open problem.

Next, we provide instructive examples through which the complete monotonicity of the SER can be seen to depend on the constellation geometry for $N^* = 3$. First, consider the constellation where the points are chosen as the vertices of a cube. In this case, the SER is given by $P_{\mathrm{e}}(\rho) = 1 - (1 - \mathcal{Q}\left(\sqrt{2\rho}\right))^3$, which can be rewritten in the form (2.2), with $\mu(u) = (3/\pi)I[1 \leq u \leq 2] + (\pi - \cos^{-1}(\alpha(u)))/2\pi^2 I[3 \leq u \leq 4] + (\pi + \cos^{-1}(\alpha(u)))/2\pi^2 I[u \geq 4]$, where $\alpha(u) = (3u^2 - 12u + 8)/(u-2)^3$. Since $\mu(u) \geq 0$, from Bernstein's theorem, $P_{\mathrm{e}}(\rho)$ is c.m.. On the other hand, consider the 3-D square QAM constellation, whose points are given by all possible sign permutations of $(\pm 1/\sqrt{6}, \pm 1/\sqrt{6}, \pm 1/\sqrt{6})$ and $(\pm 1/\sqrt{2}, \pm 1/\sqrt{2}, \pm 1/\sqrt{2})$. For this case, under the assumption of equal prior probabilities, numerical evaluation of the SER shows non-convexity (Fig. 4.3). As a result, from (2.1), the SER is not c.m.. Therefore, the c.m. properties of the SER of a constellation with $N^* > 2$ depends on the geometry and prior probabilities. With reference to existing literature, Corollary 4.3.1 is a useful generalization of [14, Theorem 2], which does not address complete monotonicity, or the possibility of reduced dimension. Further, from Corollary 4.3.1, it follows that the SER of any two-dimensional constellation under AWGN is convex. This particular consequence of Corollary 4.3.1 has been previously established in [14, Theorem 1] using a different approach. In what follows, the behavior of second derivative of the SER for constellations with $N^* > 2$ is studied.

Corollary 4.3.2. *If the reduced dimension N^* of a constellation $\mathcal{S}$ is greater than two, then the SER of the detector* (4.2) *under AWGN satisfies $P_e^{''}(\rho) \geq 0$ when $\rho \geq \rho_0$,*

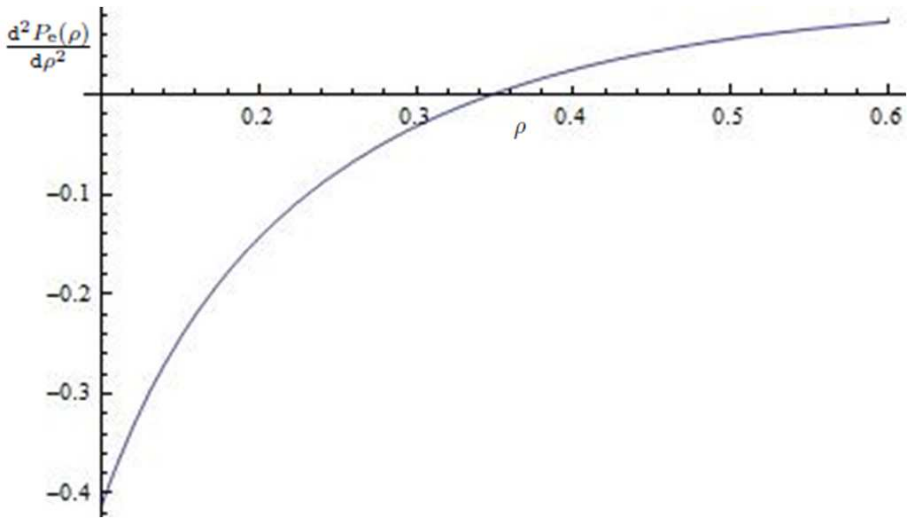

Figure 4.3: Second derivative of $P_e(\rho)$ for the 3-dimensional square QAM constellation. Since the SER is not convex, this is an example of a constellation with $N^* = 3$ whose SER is not c.m..

where $\rho_0 := 4(p + \sqrt{p})/d_{min}^2$, $p = N^*/2 - 1$, *and* d_{min} *is the minimum distance of the constellation.*

Proof. Recall from Theorem 4.3.1 that for a constellation with reduced dimension N^*, the SER can be written as

$$\overline{P}_e^{=} (\rho)^p \int_0^\infty \exp(-\rho u)\mu(u)\mathtt{d}u \,, \tag{4.20}$$

where $p = N^*/2 - 1$, with $\mu(u)$ being zero when $u \in [0, d_{\min}^2/4)$, and non-negative otherwise. We now obtain sufficient conditions for $P_e^{''}(\rho) \geq 0$. Differentiating (4.20) twice with respect to ρ, we obtain

$$P_e^{''}(\rho) = \rho^{p-2} \int_0^\infty e^{-\rho u}\mu(u) \left(u - \frac{p - \sqrt{p}}{\rho}\right)\left(u - \frac{p + \sqrt{p}}{\rho}\right)\mathtt{d}u \,, \tag{4.21}$$

where the differentiation under the integral sign in (4.20) is permitted, because the limits of integration are independent of ρ. A sufficient condition for (4.21) to be non-negative is that the integrand in (4.21) is non-negative. Accordingly, observe that the integrand in (4.21) is non-negative when $u \geq (p + \sqrt{p})/\rho$. Thus, if the support

of μ is a subset of $[(p+\sqrt{p})/\rho, \infty)$, it follows that $P_{\mathrm{e}}^{''}(\rho) \geq 0$. In other words, if $\rho \geq 4(p+\sqrt{p})/d_{\min}^2$, we have $P_{\mathrm{e}}^{''}(\rho) \geq 0$, which proves the result. $\square$

A similar but weaker result has been obtained in [14], where $\rho_0 = 4(N+\sqrt{N})/d_{\min}^2$. Corollary 4.3.2 provides a larger region where the SER has positive second derivative, which is an improvement over [14], since $N^*/2 - 1 < N^* \leq N$. Although Corollary 4.3.2 establishes a bound on the values of SNR for which the second derivative becomes non-negative, it does not forbid $P_{\mathrm{e}}^{''}(\rho) \geq 0$ for $\rho < \rho_0$. Indeed, it is possible for some multi-dimensional constellation with $N^* > 2$ to posses a convex SER. An example of such a constellation is one with $N = 3$, wherein the points correspond to the vertices of regular convex polytope (RCP) [44], for which convexity follows from the expression for the SER given by [44, Eqn. (2) - (7)].

It is useful to observe that Theorem 4.3.1 and its corollaries can be extended to the case where $\mathcal{S}$ is an arbitrary complex constellation, and the additive noise is circularly symmetric complex Gaussian noise, as shown in [50].

4.3.2 Extension to Compound Gaussian Noise

In what follows, Theorem 4.3.1 and its corollaries are generalized to the case of additive compound Gaussian noise.

The system model considered is still as in (4.1), except that the additive noise is assumed to be $\mathbf{z} = \sqrt{W}\mathbf{g}$, where W is a positive RV, which is independent of each component of $\mathbf{g} := [G_1, \ldots, G_N]^{\mathrm{T}}$, and $G_k \sim \mathcal{N}(0, 1/\rho)$ are i.i.d., for $k = 1, \ldots, N$. It should be noted that the elements of $\mathbf{z}$ are statistically dependent but uncorrelated in this case. Depending on the distribution of W, a number of noise distributions of interest arise from this formulation. For example, when W is an affine function of a Poisson RV, $\mathbf{z}$ follows a Middleton class-A distribution [15], which is used to model multi-user interference; if W is a positively skewed alpha-stable RV with a

characteristic function $\varphi_{_W}(\omega) = \exp[-|\omega|^{\alpha}(1 - j\text{sgn}(\omega)\tan\pi\alpha/2)]$, where $j = \sqrt{-1}$, $0 < \alpha < 1$, and $\text{sgn}(x)$ is the sign of x, then $\mathbf{z}$ follows the symmetric alpha-stable distribution [16], with characteristic function $\varphi_{_Z}(\omega) = \exp[-|\omega|^{2\alpha}]$. For different values of α, many impulsive noise distributions are obtained. For example, when $\alpha = 1/2$, $\mathbf{z}$ is Cauchy distributed noise [16].

In this discussion, the minimum distance detector (4.2), which is still the maximum likelihood detector for dependent but uncorrelated compound Gaussian noise, is assumed to be used at the receiver side. The reduced constellation corresponding to $\mathcal{S}$ is defined as in Definition 4.2.1 and is denoted by $\mathcal{S}^*$. It is not difficult to show that the SER of $\mathcal{S}$ and $\mathcal{S}^*$ are identical under additive compound Gaussian noise.

An extension of Theorem 4.3.1 to the case of additive compound Gaussian noise is now developed. Conditioning on $W = w$, $\mathbf{z}$ is an i.i.d. multivariate Gaussian, in which case Theorem 4.3.1 can be invoked to get $P_{\text{e}}(\rho|W = w) = \rho^p f_{\text{cm}}(\rho; w)$, where $f_{\text{cm}}(\rho; w)$ is c.m. in ρ for each w. Averaging over the distribution of W, the equivalent of Theorem 4.3.1 for the case of compound Gaussian noise is obtained, since the expectation can be interpreted as a positive linear combination of $f_{\text{cm}}(\rho; w)$, which results in a c.m. function, denoted by $f_{\text{cm}}(\rho)$.

Extensions of the corollaries of Theorem 4.3.1 for the case of compound Gaussian noise are also seen to be true, since they are obtained from the generalization of Theorem 4.3.1 to this noise model, without any additional assumptions.

4.4 Applications

In this section, applications of Theorem 4.3.1 and Corollary 4.3.1 in the context of ordering of wireless system performance is presented.

4.4.1 Applications in Stochastic Ordering

Complete monotonicity of SER for complex constellations with a reduced dimension of one or two, as suggested by extension of Corollary 4.3.1 to complex constellations finds immediate application in comparing the average SER of such constellations over two different fading channels using the theory of stochastic ordering mentioned in Section 2.3. To elucidate further, the following system model is considered:

$$\mathbf{y} = h\mathbf{s} + \mathbf{v} \,, \tag{4.22}$$

where the effect of quasi-static fading is captured by the complex scalar RV h whose real and imaginary parts are independent of each other, $\mathbf{s} \in \mathcal{S}$ with $\mathbf{s} \in \mathcal{C}^N$, and $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, (1/\rho)\mathbf{I})$ is the circularly symmetric AWGN. For this system, the instantaneous channel gain is defined as $X := |h|^2$, and the instantaneous SNR is given by ρX. Assuming that the receiver has full channel state information, the instantaneous SER can be shown to be a function of the instantaneous SNR only, and the average SER is obtained by taking the expectation of the instantaneous SER over the distribution of the instantaneous channel gain. In this application, goal is to compare the average SER of the system (4.22) under two different fading channels with instantaneous channel gains X_1 and X_2.

To begin with, let $\mathcal{S}$ be a complex constellation with a reduced dimension less than or equal to two, which has c.m. SER according to the complex extension of Corollary 4.3.1. Now, consider two fading scenarios with instantaneous channel gains X_1 and X_2, such that $X_1 \leq_{\mathrm{Lt}} X_2$. Then, according to (2.20), $\mathbb{E}\left[P_{\mathrm{e}}(\rho X_2)\right] \leq \mathbb{E}\left[P_{\mathrm{e}}(\rho X_1)\right], \forall \rho > 0$. As a result, complete monotonicity of SER can be exploited to compare two fading channels at all SNR, based on the average SER, even in cases where the expression for the average SER is not analytically tractable. As an illustrative example, consider the quadrature-PSK (QPSK) constellation, for which

the reduced dimension can be seen to be equal to 2, and $P_{\mathrm{e}}(\rho x) = \mathcal{Q}\left(\sqrt{2\rho x}\right)$ (assuming equal prior probabilities), which is c.m. in x. Now, assume QPSK is used over two different Nakagami-m fading channels, the first one with LoS parameter m_1 and instantaneous channel gain X_1, and the second one with LoS parameter m_2 and instantaneous channel gain X_2, where $m_2 \geq m_1$ so that $X_1 \leq_{\mathrm{Lt}} X_2$. In this case, (2.20) implies that $\mathbb{E}\left[\mathcal{Q}\left(\sqrt{2\rho X_2}\right)\right] \leq \mathbb{E}\left[\mathcal{Q}\left(\sqrt{2\rho X_1}\right)\right]$, $\forall \rho > 0$, which provides a way of comparing the average SERs over two fading channels with different LoS parameters.

The complex extension of Corollary 4.3.1 suggests that the SER of a complex constellation with reduced dimension greater than or equal to four is not c.m., and thus the LT ordering of instantaneous channel gains of two fading channels does not provide a conclusive comparison of the average SER of these channels. Motivated by this, a new stochastic order is introduced next, which can be used to compare the average SER of a multidimensional complex constellation over two different complex fading channels. The stochastic order $\leq_{\mathcal{G}(p)}$ is formally defined below.

Definition 4.4.1. *Let X_1 and X_2 be two positive RVs, and let $p \geq 0$ be fixed. Then $X_1 \leq_{\mathcal{G}(p)} X_2$ if and only if $\mathbb{E}\left[X_1^p \exp(-\rho X_1)\right] \geq \mathbb{E}\left[X_2^p \exp(-\rho X_2)\right]$ for all $\rho > 0$.*

In other words, for every $p \geq 0$, $\leq_{\mathcal{G}(p)}$ is an integral stochastic order in the sense of (2.13), with $\mathcal{G}p = \{g(x) \mid g(x) = -x^p \exp(-\rho x), \rho \geq 0\}$. A necessary and sufficient condition for $X_1 \leq_{\mathcal{G}(p)} X_2$ can be proved to be as follows:

Theorem 4.4.1. *Let X_1 and X_2 be two positive RVs, and $p \geq 0$. Then, $X_1 \leq_{\mathcal{G}(p)} X_2$ if and only if*

$$\mathbb{E}\left[X_1^p f_{cm}(X_1)\right] \geq \mathbb{E}\left[X_2^p f_{cm}(X_2)\right] , \tag{4.23}$$

where $f_{cm}(\cdot)$ is c.m..

Proof. For any two non-negative RVs X_1 and X_2, we have

$$X_1 \leq_{\mathcal{G}(p)} X_2 \Leftrightarrow \mathbb{E}\left[X_1^p \exp(-\rho X_1)\right] \geq \mathbb{E}\left[X_2^p \exp(-\rho X_2)\right], \forall \rho \geq 0. \tag{4.24}$$

We now establish the following:

$$X_1 \leq_{\mathcal{G}(p)} X_2 \Leftrightarrow \mathbb{E}\left[X_1^p f_{\rm cm}(\rho X_1)\right] \geq \mathbb{E}\left[X_1^p f_{\rm cm}(\rho X_1)\right] \ \forall \rho \geq 0, \tag{4.25}$$

where $f_{\rm cm}(\rho) := \int_0^\infty \exp(-\rho u)\mu(u)\mathtt{d}u$ is a c.m. function, for some $\mu(u) \geq 0$. Assume $u > 0$, then $X_1 \leq_{\mathcal{G}(p)} X_2 \Rightarrow \mathbb{E}\left[X_1^p \exp(-\rho X_1 u)\right] \geq \mathbb{E}\left[X_2^p \exp(-\rho X_2 u)\right] \ \forall \rho > 0$. Next, observe that

$$\mathbb{E}\left[X_1^p f_{\rm cm}(\rho X_1)\right] = \mathbb{E}\left[\int_0^\infty X_1^p \exp(-u\rho X_1)\mu(u)\mathtt{d}u\right] = \int_0^\infty \mathbb{E}\left[X_1^p f_{\rm cm}(\rho X_1)\right]\mu(u)\mathtt{d}u \tag{4.26}$$

$$\geq \int_0^\infty \mathbb{E}\left[X_2^p f_{\rm cm}(\rho X_2)\right]\mu(u)\mathtt{d}u = \mathbb{E}\left[X_2^p f_{\rm cm}(\rho X_2)\right], \tag{4.27}$$

$\forall \rho \geq 0$, provided the expectations exist. This proves the direct part of the theorem. To see the converse, let $f_{\rm cm}(\rho) = \exp(-\rho x)$, which is c.m. in ρ for each $x > 0$. □

Verifying the $\leq_{\mathcal{G}(p)}$ order for any pair of random variables, when $p \in \mathcal{N} \cup \{0\}$ is relatively straightforward, and can be done by comparing the p^{th} derivative of the real-valued Laplace transforms of the densities of the two RVs. Clearly, $\leq_{\mathcal{G}(p)}$ is the LT order, when $p = 0$. In this case, the envelope fading distributions for $\sqrt{X_1}$ and $\sqrt{X_2}$ such as Nakagami-m satisfy $X_1 \leq_{\mathcal{G}(0)} X_2$, when $m_1 \leq m_2$ [39]. Intriguingly however, for any $p > 0$, fading channels modelled using Nakagami-m or Rician distributions do not satisfy the $\leq_{\mathcal{G}(p)}$ order with respect to their corresponding line of sight parameters. For example, in the Nakagami-m fading scenario, $X = |h|^2$ in (4.22) is Gamma distributed. In this case,

$$\mathbb{E}\left[X^p \exp(-\rho X)\right] = \frac{m^m (m+\rho)^{-m-p}\Gamma[m+p]}{\Gamma[m]}, \tag{4.28}$$

which increases with m for small ρ and decreases otherwise, for any fixed $p > 0$. Thus, if $p > 0$, $X_1 \not\leq_{\mathcal{G}_\mathrm{p}} X_2$.

Some implications of Theorem 4.4.1 to the ordering of average SERs of multidimensional constellations over fading channels are now considered. If X_1 and X_2 are the instantaneous channel gains of two fading scenarios characterized by (4.22), then according to Theorem 4.4.1, it is easy to show that

$$X_1 \leq_{\mathcal{G}(p)} X_2 \Rightarrow \mathbb{E}\left[P_\mathrm{e}\ (\rho X_1)\right] \geq \mathbb{E}\left[P_\mathrm{e}\ (\rho X_2)\right]\ , \forall \rho > 0\ , \tag{4.29}$$

where $P_\mathrm{e}\ (\cdot)$ is the instantaneous SER of a complex constellation $\mathcal{S}$ with reduced dimension N^*, and $p = N^*/2 - 1$. This is because, from the complex extension of Theorem 4.3.1 we have $P_\mathrm{e}(\rho) = \rho^p f_\mathrm{cm}(\rho)$, which implies $P_\mathrm{e}(\rho) \in \{g(\rho) | g(\rho) = \rho^p f_\mathrm{cm}(\rho), p \geq 0\ \}$, and thus (4.29) follows from Theorem 4.4.1.

It has been reported in the literature that, it is possible to find a fading distribution such that the SER of the AWGN channel is worse than that under the fading case at low SNR, when higher dimensional constellations are employed [14]. However, examples of such fading distributions have not been a subject of investigation. Using (4.29), it is now shown that the Nakagami-m fading case is an such an example. To begin with, consider the pure AWGN channel (i.e. the no fading scenario, where $\Pr[X_1 = 1] = 1$), for which $\mathbb{E}\left[X_1^p \exp(-\rho X_1)\right] = \exp(-\rho)$. If X_2 denotes the instantaneous channel gain of a Nakagami-m fading channel, then for every $0 < m < \infty$, we now argue that there exists a $\rho_1 > 0$ such that $\mathbb{E}\left[P_\mathrm{e}(\rho X_1)\right] \geq \mathbb{E}\left[P_\mathrm{e}(\rho X_2)\right]$ for $\rho \leq \rho_1$, while $\mathbb{E}\left[P_\mathrm{e}(\rho X_1)\right] \leq \mathbb{E}\left[P_\mathrm{e}(\rho X_2)\right]$, for $\rho \geq \rho_1$. To this end, observe that $\mathbb{E}\left[X_1^p \exp(-\rho X_1)\right]$ is greater than $\mathbb{E}\left[X_2^p \exp(-\rho X_2)\right]$ when $\rho \leq \rho_1$, and vice-versa when $\rho \geq \rho_1$. Therefore, from (4.29), the AWGN channel is worse than a Nakagami-m channel in terms of SER of constellations with $N^* > 2$ at low SNR.

Next, a relation between $X_1 \leq_{\mathcal{G}q} X_2$ and $X_1 \leq_{\mathcal{G}(p)} X_2$ is obtained, where $p, q \geq 0$.

Theorem 4.4.2. *Let X_1 and X_2 be two positive RVs. Then, for $0 \leq q \leq p$, $X_1 \leq_{\mathcal{G}(p)} X_2 \Rightarrow X_1 \leq_{\mathcal{G}(q)} X_2$.*

Proof. Since $X_1 \leq_{\mathcal{G}(p)} X_2$, from Theorem 4.4.1, $\mathbb{E}\left[X_1^p f_{\text{cm}}(X_1)\right] \geq \mathbb{E}\left[X_2^p f_{\text{cm}}(X_2)\right]$, for every c.m. function $f_{\text{cm}}(\cdot)$. Choose $f_{\text{cm}}(x) := x^{-k} g_{\text{cm}}(x)$, where $0 \leq k \leq p$, and $g_{\text{cm}}(x)$ as some c.m. function. Clearly, $f_{\text{cm}}(x)$ as defined is c.m., since x^{-k} is c.m. for $k \geq 0$ and a product of c.m. functions is also c.m.. As a result, according to Theorem 4.4.1, $X_1 \leq_{\mathcal{G}(p)} X_2$ implies $\mathbb{E}\left[X_1^{p-k} g_{\text{cm}}(X_1)\right] \geq \mathbb{E}\left[X_2^{p-k} g_{\text{cm}}(X_2)\right]$. The theorem then follows by assuming $q = p - k$. □

Theorem 4.4.2 in conjunction with (4.29) implies that if $X_1 \leq_{\mathcal{G}(p)} X_2$, then X_2 is better than X_1 in terms of average SERs of all constellations at all average SNR, with the reduced dimension of the constellation satisfying $N^*/2 - 1 \leq p$. It is interesting to investigate the conditions on X_1 and X_2 such that $X_1 \leq_{\mathcal{G}(p)} X_2$ for all $p \geq 0$. In that case, X_2 will be better than X_1 in terms of average SERs of any multi-dimensional constellation at all SNRs. However, this condition is not satisfied by any pair of random variables, as described in the following Theorem:

Theorem 4.4.3. *There are no two positive random variables which satisfy $X_1 \leq_{\mathcal{G}(p)} X_2$, for all $p \geq 0$.*

Proof. Let $X_1,\ X_2$ be non-negative RVs with PDFs $f_{X_1}(x)$ and $f_{X_2}(x)$ respectively. We now show that $X_1 \leq_{\mathcal{G}(p)} X_2$ cannot hold for all $p \geq 0$, by showing that $X_1 \leq_{\mathcal{G}(p)} X_2$ does not hold for every p in a subset or $\mathbb{R} \cup \{0\}$, i.e., for $p \in \mathcal{N} \cup \{0\}$. In order to satisfy Theorem 4.4.1 for every $p \in \mathcal{N} \cup \{0\}$, we require for every p

$$(-1)^p \frac{d^p}{d\rho^p} \left[\mathbb{E}\left[\exp(-\rho X_1)\right] - \mathbb{E}\left[\exp(-\rho X_2)\right]\right] \geq 0 \ \forall \rho, \tag{4.30}$$

since $\int_0^\infty s^p \exp(-\rho s)\mu(s)\mathtt{d}s = (-1)^p(\partial^p/\partial\rho^p)\int_0^\infty \exp(-\rho s)\mu(s)\mathtt{d}s$. Recalling the definition of a c.m. function from (2.1), we gather that $\mathbb{E}\left[\exp(-\rho X_1)\right] - \mathbb{E}\left[\exp(-\rho X_2)\right]$

in (4.30) must be a c.m. function. Consequently, Bernstein's Theorem mandates that $f_{X_1}(x) - f_{X_2}(x) \geq 0 \; \forall x$. However, this condition is never satisfied by any pair of random variables, since both the density functions must individually integrate to unity, which cannot be the case if $f_{X_1}(x) \geq f_{X_2}(x) \; \forall x$. Thus, the Theorem follows. $\square$

As described in the paragraph above Theorem 4.4.2, the AWGN channel is not the best channel in terms of SER of constellations with $N^* > 2$ at all SNR. According to Theorem 4.4.3, there is no fading distribution which dominates every other fading distribution in the sense of $\leq_{\mathcal{G}(p)}$ for all $p \geq 0$. Therefore, unlike cases where the SER metric is convex, where AWGN (no fading) outperforms any fading, this is not the case for $N^* > 2$. Moreover, Theorem 4.4.3 suggests that there is no fading distribution which serves the role of "best" fading distribution, in terms of SERs of constellations of every dimension.

Chapter 5

ERGODIC CAPACITY ORDERING USING THE SHANNON TRANSFORM

5.1 Motivation and Literature Survey

In this chapter, a stochastic order which can be used to compare fading channels based on the Shannon transform of the instantaneous SNR is discussed. A fading channel is said to be better than another in the ergodic capacity order, if its corresponding Shannon transform is bigger for all ρ. The proposed order is a kind of stochastic order on positive RVs. Previously, the stochastic LT order, which compares the real-valued Laplace transforms of RVs has been used to compare two fading distributions and applied to comparing the average error rate of M-QAM modulations (See Chapter 3 or [39]). This can be explained by the fact that error rates of some modulations are non-negative integral mixtures of decaying exponentials, which can also be viewed as the Laplace transform. It has been shown in Chapter 3 that LT ordering of instantaneous SNRs implies ordering of ergodic capacities, but not conversely.

The ergodic capacity order presented in this Chapter is new to both stochastic ordering literature as well as information theory literature. Many parametric fading distribution families such as the Nakagami-m, Rician and Hoyt are observed to have the property that the ergodic capacity is monotone with respect to the line of sight (LoS) parameter for each of these distributions, as shown herein. Consequently, the instantaneous SNR of these fading channels serve as examples of ergodic capacity ordered random variables. The properties of this stochastic order are useful in obtaining comparisons of the performance of systems involving multiple SNR RVs, as

described in Section 5.3. For example, let $\{X_i\}_{i=1}^M$ and $\{Y_i\}_{i=1}^M$ be two sets of fading channels such that the ergodic capacity of X_i is less than that of Y_i, $i = 1, \ldots, M$ at all SNR. Then, the properties of the ergodic capacity order provide the conditions under which a composite system consisting of $\{X_i\}_{i=1}^M$ as the component fading channels has a smaller ergodic capacity than that of a system with components $\{Y_i\}_{i=1}^M$. Such comparisons of ergodic capacities can be made even in cases when a closed-form expression is not available. A MIMO extension of the definition of the ergodic capacity order, which can be used to order positive semidefinite symmetric random matrices is given in Section 5.4.

We now proceed to define a stochastic order for comparing fading distributions based on the Shannon transform, which is defined in Section 2.2.2.

5.2 The Ergodic Capacity Order

Recall that the ergodic capacity of a single-input single-output (SISO) system is given by $\mathbb{E}\left[\log\left(1+\rho X\right)\right]$, where X is the square of the amplitude of the complex fading gain, and is defined as the instantaneous fading power of the channel. It is straightforward to see through an application of Jensen's inequality that the AWGN channel (with no fading) outperforms every fading distribution with same average channel power, in terms of the ergodic capacity at all SNR. However, given two fading distributions, it is not trivial to compare them based on the ergodic capacity, as obtaining a closed-form expression for the ergodic capacity of many fading channels is analytically intractable. Motivated by this, we propose a stochastic ordering method, which can be used to compare the ergodic capacity of two different fading channels.

5.2.1 Definition

Definition 5.2.1. *If X and Y are arbitrary nonnegative RVs, then X is said to be dominated by Y in the ergodic capacity order (i.e. $X \leq_c Y$), if the Shannon transforms of X and Y exist and $\overline{C}^{(X)}(\rho) \leq \overline{C}^{(Y)}(\rho)$ for $\rho \geq 0$.*

For this stochastic order, the generator is chosen as

$$\mathcal{G} = \{g(x) : g(x) = \log(1 + \rho x), \rho \geq 0\}. \tag{5.1}$$

Distributions of interest for which the ergodic capacity is finite at all finite SNR can be determined using either Proposition 2.2.1 or Proposition 2.2.2. Next, some useful properties of the capacity order and a few examples of ergodic capacity ordered RVs are discussed.

5.2.2 Properties

The following properties hold for nonnegative RVs.

S1: $X \leq_c Y \iff \mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)]$, $\forall g \in \mathcal{TBF}$, such that the expectations exist.

S2: $X \leq_c Y \iff g(X) \leq_c g(Y)$, $\forall g \in \mathcal{CTBF}$.

S3: $X \leq_{\text{Lt}} Y \implies X \leq_c Y$.

S4: Let $X_1, \ldots, X_M$ be independent and $Y_1, \ldots, Y_M$ be independent. If $X_m \leq_c Y_m$, $m = 1, \ldots, M$, then $g(X_1, \ldots, X_M) \leq_c g(Y_1, \ldots, Y_M)$, $\forall g \in \mathcal{CTBF}_M$.

S5: If $X \leq_c Y$ and $Y \leq_c Z$, then $X \leq_c Z$.

S6: If $X \leq_c Y$ and $Y \leq_c X$, then $F_X(\cdot) = F_Y(\cdot)$ a.e.

The proofs of these properties can be obtained from the proofs of the more general MIMO ergodic capacity order (introduced in Section 5.4), which are presented at the end of this Chapter. A straightforward implication of Property S1 is that if $X \leq_{\mathrm{c}} Y$, then $\mathbb{E}[X] \leq \mathbb{E}[Y]$, since $g(x) = x$ is a Thorin-Bernstein function. In other words, if one fading channel has a higher ergodic capacity than another at all SNR, then it is necessary that the average fading power of the first channel is no smaller than that of the second. Properties S5 and S6 together constitute the definition of a partial order, and consequently $\leq_{\mathrm{c}}$ is a partial order on nonnegative RVs.

Interpreting ρX and ρY as the instantaneous SNRs of two different fading channels, Properties S1-S6 are useful in obtaining the conditions under which the ergodic capacity of a composite system with coding/decoding capabilities only at the transmitter/receiver under the channel Y is greater than that under X at all SNR. Although Property S3 suggests that every pair of Laplace transform ordered random variables also obey the ergodic capacity order, the converse is not true in general. A counterexample can be found in [39, 51]. Thus, it is possible that the average symbol error rate of differential binary phase shift keying modulation in channel X is less than that in Y at high SNR, while the situation reverses when the capacity achieving code is applied on both channels. Interpreting the ergodic capacity as what is achievable by coding over an i.i.d. time-extension of the channel, we reach the conclusion that even though Y offers more diversity than X for an uncoded system, the i.i.d. extension of X lends itself to more diversity than that of Y. To put it more simply, at high SNR, it is possible for one fading channel to be superior to another in terms of error rates in the absence of coding, while being inferior when the capacity achieving code is employed over both channels.

5.2.3 Extension to Nonnegative Measures

Shannon Transform Order

In what follows, a generalized version of the ergodic capacity order applicable to order nonnegative measures is described.

Let μ be a non-negative measure on $\mathbb{R}^+$. Its Shannon transform is defined as $\int_0^\infty \log(1+s/u)\mu(\mathtt{d}u)$. If μ_1 and μ_2 are two non-negative measures for which the Shannon transforms exist and are finite, then we say μ_1 is dominated by μ_2 in the Shannon transform sense, and write $\mu_1 \leq_{\mathrm{S}} \mu_2$ to denote $\int_0^\infty \log(1+s/u)\mu_1(\mathtt{d}u) \leq \int_0^\infty \log(1+s/u)\mu_2(\mathtt{d}u)$, for all $s \geq 0$. Note that when μ_1 and μ_2 are probability measures, then the ergodic capacity order defined in Definition 5.2.1 can be obtained.

Some properties of the Shannon transform order are the following:

P1: If $\int_0^\infty \mu_1(\mathtt{d}u) = \int_0^\infty \mu_2(\mathtt{d}u) < \infty$, then $\mu_1 \leq_{\mathrm{Lt}} \mu_2 \implies \mu_1 \leq_{\mathrm{S}} \mu_2$

P2: $\mu_1 \leq_{\mathrm{S}} \mu_2 \iff \int_0^\infty \sigma(u)\mu_1(\mathtt{d}u) \leq \int_0^\infty \sigma(u)\mu_2(\mathtt{d}u)$, for all $\{\sigma \in \mathcal{TBF} | \sigma(0) = 0\}$.

P3: Let $\int_0^\infty \mu_1(\mathtt{d}u) \leq \int_0^\infty \mu_2(\mathtt{d}u)$. Then $\mu_1 \leq_{\mathrm{S}} \mu_2 \iff \int_0^\infty \sigma(u)\mu_1(\mathtt{d}u) \leq \int_0^\infty \sigma(u)\mu_2(\mathtt{d}u)$, for all $\sigma \in \mathcal{TBF}$.

Applications of this generalization can be found in Chapter 5.6.

5.2.4 Examples

Next, we give examples of pairs of RVs X, Y relevant to wireless communications, for which $X \leq_{\mathrm{c}} Y$ holds. In general, establishing ergodic capacity ordering using its definition is often inconclusive, since the corresponding integrals are intractable. Fortunately, using Property S3, it is possible to provide examples of pairs of RVs which obey capacity ordering. In what follows, examples of parametric fading distributions which obey the ergodic capacity order are given. These distributions are also known to satisfy the Laplace transform order [39].

Nakagami Fading

The Nakagami-m fading model, for which the envelope $\sqrt{X}$ is Nakagami distributed, and the instantaneous fading power X is Gamma distributed, with PDF given by

$$f_X(x) = \frac{m^m}{\Gamma(m)} x^{m-1} \exp(-mx)\ , x \geq 0\ , \tag{5.2}$$

where $m > 0$ is the line of sight parameter, and $\Gamma(r) := \int_0^\infty t^{r-1} \exp(-t) \mathtt{d}t$ is the gamma function. Let $X \sim \text{Gamma}(m^X)$, and $Y \sim \text{Gamma}(m^Y)$ with $m^X \leq m^Y$. For this case, it is easy to verify that $X \leq_{\text{Lt}} Y$, which implies that $X \leq_{\text{c}} Y$, according to Property S3. Property S3 requires the existence of the Shannon transforms, which is proved as follows. Observing that $\mathbb{E}[X] = \mathbb{E}[Y] = 1$ is finite, from Proposition 2.2.2, the Shannon transforms exist. This is because setting $\delta = 1$ in Proposition 2.2.2 is equivalent to saying that the mean value is finite.

Rician Fading

The Rician fading model: In this case, the envelope of the fading i.e., $\sqrt{X}$ is Rice distributed with line of sight parameter K, and the corresponding instantaneous fading power distribution is given by

$$f_X(x) = (K+1) \exp\left[-(K+1)x - K\right] I_0\left(2\sqrt{K(K+1)x}\right)\ , \tag{5.3}$$

where $I_0(t) := \sum_{m=0}^{\infty} (t/2)^{2m}/(m!\Gamma(m+1))$ is the modified Bessel function of the first kind of order zero. If the distribution of X and Y have parameters K^X and K^Y respectively, with $K^X \leq K^Y$, then $X \leq_{\text{c}} Y$. The existence of the Shannon transforms is established in way similar to that of the Nakagami-m case.

Hoyt Fading

The Nakagami-q (Hoyt) fading model: The envelope $\sqrt{X}$ of the fading RV is Hoyt distributed, and the density of the instantaneous fading power is given by

$$f_X(x) = a \exp(-a^2 x) I_0(bx) , \tag{5.4}$$

where $a = (1+q^2)/2q$, $b = (1-q^4)/4q^2$. If X and Y have parameters q^X and q^Y respectively, where $q^X \leq q^Y$, then $X \leq_{\mathrm{c}} Y$. The existence of the Shannon transforms is established in way similar to that of the Nakagami-m case.

For the cases of Nakagami, Rician and Hoyt fading, the increase in ergodic capacity with increase in the LoS parameter of the distribution is not due to an increase in the average fading power, since $\mathbb{E}[X] = \mathbb{E}[Y]$, which is independent of the LoS parameter.

In what follows, we show that ergodic capacity ordering of a given SISO system under two different fading channels can be used to make meaningful conclusions when a number of such systems are combined to form a system involving multiple random variables.

5.3 Systems Involving Multiple Random Variables

In order to illustrate the applicability of the ergodic capacity order to compare the performance of systems, we provide examples of composite systems where ergodic capacity ordering of component SISO systems can be used to conclude the capacity ordering of the system, and also some applications where this is not necessarily the case. Such generic conclusions can be made even when closed form expressions for the ergodic capacity are not available. Throughout, we assume that the receiver has a perfect estimate of the instantaneous fading power, while the transmitter does not possess any such information.

5.3.1 Diversity Combining Systems

As examples of systems involving multiple fading links, we first consider diversity combining schemes such as maximum ratio combining (MRC) and equal gain combining (EGC) using M receive antennas, for which we aim to compare the ergodic capacity under two different fading scenarios. Using the properties of the ergodic capacity order, we now show that diversity combining systems formed using a better set of components yields a system with a higher ergodic capacity, for the two schemes considered.

Maximum Ratio Combining

Conditioned on the instantaneous fading power $X_m = x_m$, $m = 1, \ldots, M$, the fading power after combining is given by

$$g_{\text{MRC}}(x_1, \ldots, x_M) = \sum_{m=1}^{M} x_m \ . \tag{5.5}$$

The ergodic capacity corresponding to this combining scheme is given by

$$\overline{C}^{(X)}_{\text{MRC}}(\rho) = \mathbb{E}\left[\log\left(1 + \rho g_{\text{MRC}}(X_1, \ldots, X_M)\right)\right] \ . \tag{5.6}$$

It is easy to see that $\overline{C}^{(Y)}_{\text{MRC}}(\rho)$ is finite if the Shannon transforms of Y_m, $m = 1, \ldots, M$ exist. We then obtain the following result, which can be used to compare the ergodic capacity of MRC in two different fading environments characterized by instantaneous fading powers $(X_1, \ldots, X_M)$ and $(Y_1, \ldots, Y_M)$:

Proposition : If $X_m \leq_{\text{c}} Y_m$, $m = 1, \ldots, M$, then $\overline{C}^{(X)}_{\text{MRC}}(\rho) \leq \overline{C}^{(Y)}_{\text{MRC}}(\rho)$, at all $\rho \geq 0$.

Proof. We first verify that $g_{\text{MRC}}(\cdot)$ is a composable Thorin-Bernstein function. Then, we use Property S4 to conclude $\overline{C}^{(X)}_{\text{MRC}}(\rho) \leq \overline{C}^{(Y)}_{\text{MRC}}(\rho)$, at all $\rho \geq 0$, when $X_m \leq_{\text{c}} Y_m$, $m = 1, \ldots, M$.

To show that $g_{\text{MRC}}(\cdot) \in \mathcal{CTBF}$, treat x_1 in $g_{\text{MRC}}(\cdot)$ as the variable, while treating other arguments as constants, to get $g_{\text{MRC}}(x_1; x_2, \ldots, x_M) = x_1 + k$, where $k = \sum_{m=2}^{M} x_m$. By definition, $g_{\text{MRC}} \in \mathcal{CTBF}$ if and only if

$$h_{\text{MRC}}(x) := \frac{g'_{\text{MRC}}(x; x_2, \ldots, x_M)}{g_{\text{MRC}}(x; x_2, \ldots, x_M)} = (x + k)^{-1} \tag{5.7}$$

is a Stieltjes function. This is indeed the case, since $h_{\text{MRC}}(\cdot)$ satisfies (2.3) with $a = 0, b = 0$, and $\mu(s) = \delta(s)$. Now, assuming $X_m \leq_{\text{c}} Y_m, m = 1, \ldots, M$, we have from Property S4 $g_{\text{MRC}}(X_1, \ldots, X_M) \leq_{\text{c}} g_{\text{MRC}}(Y_1, \ldots, Y_M)$, which implies $\overline{C}^{(X)}_{\text{MRC}}(\rho) \leq \overline{C}^{(Y)}_{\text{MRC}}(\rho)$, at all $\rho \geq 0$. □

Thus, if Y_m dominates X_m in the ergodic capacity order for $m = 1, \ldots, M$, then the MRC system with fading links given by $Y_1, \ldots, Y_M$ will have a higher ergodic capacity than that with $X_1, \ldots, X_M$ at all SNR.

Equal Gain Combining

For the case of equal gain combining, the ergodic capacity is given by

$$\overline{C}^{(X)}_{\text{EGC}}(\rho) = \mathbb{E}\left[\log\left(1 + \rho g_{\text{EGC}}(X_1, \ldots, X_M)\right)\right] , \tag{5.8}$$

where $g_{\text{EGC}}(\cdot)$ represents the combined instantaneous fading power, and is given by

$$g_{\text{EGC}}(x_1, \ldots, x_M) = M^{-1}\left(\sum_{m=1}^{M} \sqrt{x_m}\right)^2 . \tag{5.9}$$

It is possible to show that $\overline{C}^{(Y)}_{\text{EGC}}(\rho)$ is finite if the Shannon transforms of Y_m, $m = 1, \ldots, M$ exist, by using the Cauchy-Schwarz inequality in addition to showing that the Shannon transform of $\sqrt{Y}_m$ exists if the Shannon transform of Y_m exists. While closed-form expressions for the ergodic capacity of equal gain combining for several fading distributions are unknown, it is still possible for us to compare these quantities using the ergodic capacity ordering of component branches:

Proposition : Let $X_m \leq_{\mathrm{c}} Y_m$, $m = 1, \ldots, M$. Then $\overline{C}^{(X)}_{\mathrm{EGC}}(\rho) \leq \overline{C}^{(Y)}_{\mathrm{EGC}}(\rho)$, at all $\rho \geq 0$.

Proof. We first prove that $g_{\mathrm{EGC}} \in \mathcal{CTBF}$, and then use Property S4 to complete the proof. In order to show that $g_{\mathrm{EGC}} \in \mathcal{CTBF}$, treat x_1 as the variable and all the other arguments of g_{EGC} as constants, so that $g_{\mathrm{EGC}}(x_1; x_2, \ldots, x_M) = M^{-1}(x_1 + 2\sqrt{x_1}k + k^2)$, where $k = \sum_{m=2}^{M} x_m$. By definition, $g_{\mathrm{EGC}}(\cdot)$ in $\mathcal{CTBF}$ if and only

$$h(x) := \frac{g'_{\mathrm{EGC}}(x; k)}{g_{\mathrm{EGC}}(x; k)} = (x + k\sqrt{x})^{-1} \tag{5.10}$$

is a Stieltjes function. To show that $h \in \mathcal{S}$, observe that $(h(x^{-1}))^{-1} = x^{-1} + kx^{-1/2}$ is a Stieltjes function, since any function of the form $x^{\alpha-1}, 0 \leq \alpha \leq 1$ is a Stieltjes function [17, p. 13], and positive linear combinations of Stieltjes functions also yields a Stieltjes function. To complete the argument, since $(h(x^{-1}))^{-1} \in \mathcal{S}$, $h(x)$ must also belong to $\mathcal{S}$ [17, p. 66]. Consequently, $g_{\mathrm{EGC}}(\cdot) \in \mathcal{CTBF}$. The rest of the proof follows arguments similar to the MRC case. □

Using Proposition 5.3.1, we infer that if a collection of SISO systems with higher ergodic capacity is combined to form an EGC system, then the composite EGC system will have higher overall ergodic capacity.

5.3.2 Multi-Hop Amplify-Forward Relay System

We now turn our attention to multi-hop amplify-forward (MH-AF) relay systems. This is an example of a system where despite component-wise ergodic capacity ordering of individual hops, the overall system need not have a higher ergodic capacity at all SNR. The system consists of a source, which transmits data to a destination using $M - 1$ half-duplex variable gain relays, which possess receive CSI (Figure 5.1). The source transmits in time slot 1 to relay 1, and relay m in turn amplifies and

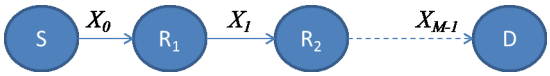

Figure 5.1: M-hop relay. S represents the source, R_m represent the relays and D represents the destination.

retransmits to relay $m+1$ in time slot $m+1$, $m = 1, \ldots, M-2$, while relay $M-1$ amplifies and transmits to the destination in time slot M. The gain of the m^{th} relay node is given by $\alpha_m = \rho/(\rho X_{m-1} + 1)$ [52], where X_m is the instantaneous fading power of the m^{th} hop, for $m = 1, \ldots, M-1$. X_0 denotes the instantaneous fading power of the channel between the source and the first relay node. It is assumed that coding/decoding capabilities are provided to the transmitter/receiver alone. In this case, the end-to-end ergodic capacity is given by

$$\overline{C}^{(X)}_{\mathrm{MH-AF}}(\rho) = \mathbb{E}\left[\log\left(1 + g_{\mathrm{MH-AF}}(X_0, \ldots, X_{M-1})\right)\right] , \tag{5.11}$$

where $g_{\mathrm{MH-AF}}(x_0, \ldots, x_{M-1}) := (\prod_{m=0}^{M-1}[(1 + (\rho x_m)^{-1})] - 1)^{-1}$. Exact expressions for the ergodic capacity in arbitrary fading channels are intractable, even for the two-hop case. Previously, the ergodic capacity of such a relay in fading channels has been obtained as an infinite series in [53]. Nevertheless, even in the absence of closed-form expressions, it is possible to compare the ergodic capacities of two such relay networks which are identical, except for the fading distribution across the hops. In order to compare the performance of the MH-AF relay in two different fading scenarios, let X_m and Y_m denote the instantaneous fading power of the m^{th} link of the first and second fading channels respectively, for $m = 0, \ldots, M-1$.

Proposition : If $X_m \leq_{\mathrm{Lt}} Y_m$, $m = 0, \ldots, M-1$, then $\overline{C}^{(X)}_{\mathrm{MH-AF}}(\rho) \leq \overline{C}^{(Y)}_{\mathrm{MH-AF}}(\rho)$ at all $\rho \geq 0$.

Proof. To establish this result, we recall that a property similar to Property S4 holds for LT ordered random variables: If $X_m \leq_{\mathrm{Lt}} Y_m, m = 0, \ldots M-1$, then

$g(X_0, \ldots, X_{M-1}) \leq_{\text{Lt}} g(Y_0, \ldots, Y_{M-1})$, whenever g which is a Bernstein function in each variable, while viewing all the other variables as constants [3, Theorem 5.A.7]. Now, this can be established by straight-forward differentiation with respect to x_i. As a result, if the instantaneous fading powers satisfy $X_m \leq_{\text{Lt}} Y_m, m = 0, \ldots, M-1$, then $g_{\text{MH-AF}}(X_0, \ldots, X_{M-1}) \leq_{\text{Lt}} g_{\text{MH-AF}}(Y_0, \ldots, Y_{M-1})$, and therefore we observe from Property S3 that

$$g_{\text{MH-AF}}(X_0, \ldots, X_{M-1}) \leq_{\text{c}} g_{\text{MH-AF}}(Y_0, \ldots, Y_{M-1}). \tag{5.12}$$

The proposition then follows, since ergodic capacity ordered RVs have ordered expectations. □

In other words, if each hop of Y dominates the corresponding hop of X in the Laplace transform order, then the overall ergodic capacity of the M-hop MH-AF relay formed using $\{Y_m\}_{m=0}^{M-1}$ will be higher than that formed using $\{X_m\}_{m=0}^{M-1}$.

However, this conclusion does not hold if we make the weaker assumption that $X_m \leq_{\text{c}} Y_m$, instead of $X_m \leq_{\text{Lt}} Y_m$, $m = 0, \ldots, M-1$. In other words, componentwise ordering of links in the ergodic capacity ordering sense does not imply the ordering of the overall system. To see a counterexample, consider the case of an interference dominated channel, where the instantaneous fading power to interference power ratio X_m are independent and Pareto-type distributed with parameter β_X [31]:

$$F_{X_m}(x) = \frac{x^{\beta_X}}{1 + x^{\beta_X}}, x > 0, \beta_X > 0\ , \tag{5.13}$$

and Y_m similarly with parameter β_Y, where $\beta_X \leq \beta_Y$. In this case, it can be shown that $X_m \leq_{\text{c}} Y_m$, but $X_m \not\leq_{\text{Lt}} Y_m$, $m = 0, \ldots, M-1$. As an illustrative example, Fig. 5.2 shows the numerically evaluated ergodic capacities of a multi-hop relay with $M = 3$ hops under Pareto-type distributed signal-to-interference ratio with parameters $\beta_X = 1$ and $\beta_Y = 3$, so that for each hop $X_m \leq_{\text{c}} Y_m, m = 0, 1, 2$ is

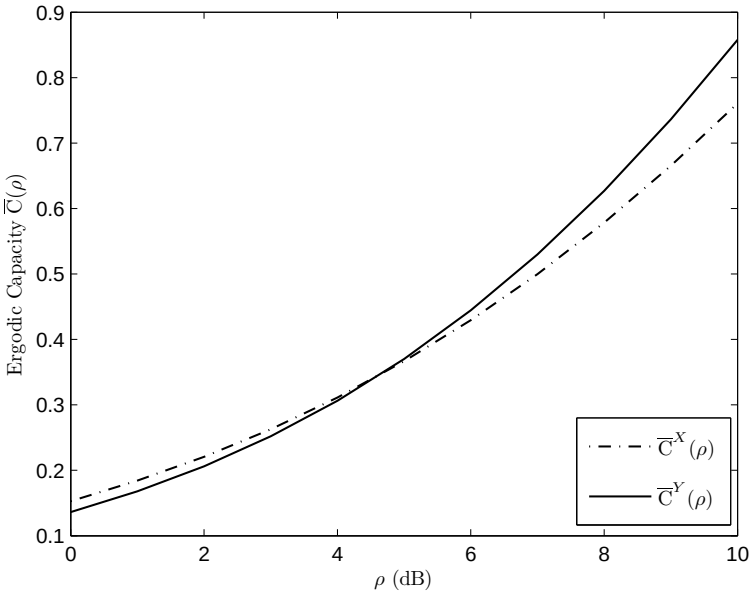

Figure 5.2: Ergodic capacity of amplify-forward relay with $M = 3$ slots. The instantaneous SINR is Pareto distributed with parameters $\beta_X = 1$ (dashed line) and $\beta_Y = 3$ (solid line).

satisfied. It is observed from Fig. 5.2 that for $\rho < \rho_0$, where $\rho_0 \approx 5$ dB, X is a better channel than Y in the ergodic capacity order, while for $\rho \geq \rho_0$, the situation is reversed. In summary, the MH-AF system is an example of a case where contrary to intuition, it is possible for a fading channel system $\{Y_m\}_{m=0}^{M-1}$ to not have a higher ergodic capacity at all SNR than that of $\{X_m\}_{m=0}^{M-1}$, even though the ergodic capacity of each Y_m is higher than that of X_m, $m = 0, \ldots, M-1$ at all SNR.

5.3.3 Fading Multiple Access Channel

In this example, we focus on comparing the ergodic capacity regions of a multi-user Gaussian MAC network in two different fading scenarios. Consider the following system model:

$$r = \sqrt{\rho} \sum_{m=1}^{M} h_m s_m + v \ , \tag{5.14}$$

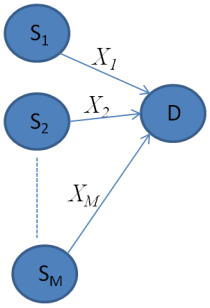

Figure 5.3: M-user multiple access channel.

where r is the received signal, ρ is the average SNR of each user, s_m is the transmitted symbol of user m, h_m is the complex i.i.d (across time) ergodic fading between each user and the destination, and v is the AWGN at the receiver. It is assumed that only the receiver possesses CSI of all the users. The receiver intends to decode the signals from all the users. If $X_m := |h_m|^2, m = 1, \ldots, M$, then the ergodic capacity region $\overline{C}^{(\cdot)}_{\text{MAC}}(\rho)$ is the set of all rate M-tuples that satisfy [54, pp. 407],

$$\sum_{m \in \mathcal{S}} R_m(\rho) \leq \mathbb{E}\left[\log\left(1 + \rho \sum_{\mathcal{S}} X_m\right)\right] , \tag{5.15}$$

where $\mathcal{S} \subset 2^{\{1,\ldots,M\}}$. Using the ergodic capacity order, we can now make the following observation which links the ordering of ergodic capacities of each user to the overall ergodic capacity region of the fading MAC.

Proposition : If $X_m \leq_{\text{c}} Y_m, m = 1, \ldots, M$, then $\overline{C}^{(X)}_{\text{MAC}}(\rho) \subseteq \overline{C}^{(Y)}_{\text{MAC}}(\rho)$, for $\rho \geq 0$.

Proof. To begin with, observe that $g_{\text{MAC},\mathcal{S}}(x_1, \ldots, x_M) := \sum_{\mathcal{S}} x_m$ belongs to $\mathcal{CTBF}_{|\mathcal{S}|}$. If $X_m \leq_{\text{c}} Y_m, m = 1, \ldots, M$, then

$$g_{\text{MAC},\mathcal{S}}(X_1, \ldots, X_M) \leq_{\text{c}} g_{\text{MAC},\mathcal{S}}(Y_1, \ldots, Y_M), \forall \mathcal{S} \subset 2^{\{1,\ldots,M\}} , \tag{5.16}$$

due to Property S4. Hence, if $X_m \leq_{\text{c}} Y_m, m = 1, \ldots, M$, then $\overline{C}^{(X)}_{\text{MAC}}(\rho) \subseteq \overline{C}^{(Y)}_{\text{MAC}}(\rho)$, for all $\rho \geq 0$. □

In other words, if each user of the system X has a higher ergodic capacity than the corresponding user in the system Y, then $\overline{C}^{(X)}_{\text{MAC}}(\rho) \subseteq \overline{C}^{(Y)}_{\text{MAC}}(\rho)$, for $\rho \geq 0$.

5.4 MIMO Ergodic Capacity Order

In this section, the ergodic capacity ordering of MIMO systems is presented. Some properties of this stochastic order are discussed, and an application of this framework in a MIMO MAC setting is presented. Before doing so, we formally define a MIMO system through its single letter characterization:

$$\mathbf{y} = \sqrt{\rho}\mathbf{H}\mathbf{s} + \mathbf{v} \ , \tag{5.17}$$

where $\mathbf{y}$ is the received signal, $\mathbf{H}$ is a complex $N_R \times N_T$ random matrix which captures the effect of ergodic quasi-static fading, $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ is the additive noise, $\mathbf{s}$ is the transmitted symbol vector, and ρ is the average SNR per transmit antenna. $\mathbf{H}$ and $\mathbf{v}$ are assumed to be i.i.d across time, as a result of which a time index has not been used in (5.17). Further, it is assumed that the receiver tracks the channel fading realizations $\mathbf{H}$, while no such CSI is available at the transmitter. For this system model, the instantaneous fading power is given by $\mathbf{H}^{\mathrm{H}}\mathbf{H}$, and is denoted as $\mathbf{X}$. In this case, the ergodic capacity is the Shannon transform of the instantaneous fading power, and is given by $\overline{C}^{(X)}_{\mathrm{MIMO}}(\rho) = \mathbb{E}\left[\log\det\left(\mathbf{I} + \rho\mathbf{X}\right)\right]$.

Remark: The Shannon transform for an arbitrary distribution on positive semidefinite matrices need not exist. Using Proposition 2.2.2, it can be shown that the Shannon transform for a positive semidefinite matrix $\mathbf{X}$ exists, if there exists some $\delta \in (0, 1]$, such that $\int_0^t \mathbb{E}\left[1 - F_{\lambda_Q(\mathbf{X})}(u)\right] \mathtt{d}u = O(t^{1-\delta}), t \to \infty$, where Q is uniformly picked from $\{1, \ldots, n\}$.

In what follows, we define a partial order on the instantaneous fading power, which can be used to compare the ergodic capacity of composite MIMO systems under two different fading environments.

5.4.1 Definition and Properties

Definition 5.4.1. *For two random positive semidefinite matrices* $\mathbf{X}$, $\mathbf{Y}$, *we say that* $\mathbf{X}$ *is dominated by* $\mathbf{Y}$ *in the MIMO ergodic capacity order, and write* $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y}$, *if the Shannon transforms of* $\mathbf{X}$ *and* $\mathbf{Y}$ *exist and* $\mathbb{E}\left[\mathbf{tr}\log\left(\mathbf{I}+\rho\mathbf{X}\right)\right] \leq \mathbb{E}\left[\mathbf{tr}\log\left(\mathbf{I}+\rho\mathbf{Y}\right)\right]$, *for all* $\rho \geq 0$.

In Definition 5.4.1, $\log(\cdot)$ is to be viewed as a matrix function, in the sense of Section 2.1.5. It is easy to show that $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y}$ is equivalent to $\mathbb{E}\left[\log\det\left(\mathbf{I}+\rho\mathbf{X}\right)\right] \leq \mathbb{E}\left[\log\det\left(\mathbf{I}+\rho\mathbf{Y}\right)\right]$, at all $\rho \geq 0$. In contrast to the ergodic capacity order on random variables, the MIMO ergodic capacity corresponding to two different random matrices $\mathbf{X}$ and $\mathbf{Y}$ may be identical (for example, when $\mathbf{Y} = \mathbf{U}\mathbf{X}\mathbf{U}^{\mathrm{H}}$, where $\mathbf{U}$ is a unitary matrix). In this circumstance, we write $\mathbf{X} =_c \mathbf{Y}$. In what follows, some properties of the MIMO ergodic capacity order are developed, which can be viewed as matrix analogues to the properties developed in Section 5.2.2. The following properties are true for random matrices $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathbb{S}_+^n$, for which the Shannon transforms exist.

M1: $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y} \iff \mathbb{E}\left[\mathbf{tr}\; g(\mathbf{X})\right] \leq \mathbb{E}\left[\mathbf{tr}\; g(\mathbf{Y})\right]$, for all $g : \mathbb{R} \to \mathbb{R}$, such that $g \in \mathcal{TBF}$, provided the expectations exist.

M2: $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y} \iff g(\mathbf{X}) \preceq_{\mathrm{c}} g(\mathbf{Y})$, for all $g : \mathbb{R} \to \mathbb{R}$, such that $g \in \mathcal{CTBF}$.

M3: $\mathbb{E}\left[\mathbf{tr}\; \exp(-\rho\mathbf{X})\right] \geq \mathbb{E}\left[\mathbf{tr}\; \exp(-\rho\mathbf{Y})\right]\ \forall \rho \geq 0$ then $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y}$.

M4: Let $\{\mathbf{X}_m\}_{m=1}^M$, $\{\mathbf{Y}_m\}_{m=1}^M$ be independent random matrices in $\mathbb{S}_+^n$, such that $\mathbf{X}_m \preceq_{\mathrm{c}} \mathbf{Y}_m$, $m = 1, \ldots, M$. Let $g(\mathbf{X}_{1:M}) := g(\mathbf{X}_1, \ldots, \mathbf{X}_M)$, i.e., g operates on M $\mathbb{S}_+^n$ matrices and produces a $\mathbb{S}_+^n$ matrix. If $g : \mathbb{R}^M \to \mathbb{R}$ is such that $g \in \mathcal{CTBF}_M$ then $g(\mathbf{X}_{1:M}) \preceq_{\mathrm{c}} g(\mathbf{Y}_{1:M})$.

M5: If $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y}$, and $\mathbf{Y} \preceq_{\mathrm{c}} \mathbf{Z}$, then $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Z}$.

M6: $\mathbf{X} =_c \mathbf{Y}$ if and only if $\sum_{i=1}^{n} F_{\lambda_i(\mathbf{X})}(u) = \sum_{i=1}^{n} F_{\lambda_i(\mathbf{Y})}(u)$, where $F_{\lambda_i(\mathbf{X})}(\cdot)$ is the marginal CDF of the i^{th} largest eigenvalue of $\mathbf{X}$.

The proofs of properties M1-M4, and M6 can be found toward the end of this Chapter, while Property M5 is straight-forward to establish, and its proof is omitted. Property M3 provides a useful sufficient condition to verify if two random matrices obey the MIMO ergodic capacity order. This is because

$$\mathbb{E}\left[\mathbf{tr}\, e^{-\rho\mathbf{X}}\right] \geq \mathbb{E}\left[\mathbf{tr}\, e^{-\rho\mathbf{Y}}\right] \iff \sum_{i=1}^{n} \mathbb{E}\left[e^{-\rho\lambda_i(\mathbf{X})}\right] \geq \sum_{i=1}^{n} \mathbb{E}\left[e^{-\rho\lambda_i(\mathbf{Y})}\right], \forall \rho \geq 0, \quad (5.18)$$

and Laplace transforms of the eigenvalue distributions are more easy to compute, when compared to the expectations of the log-determinants.

Next, we form an interesting interpretation of Property M6. From Property M6, it follows that $\mathbf{X} =_c \mathbf{Y} \iff \mathbb{E}_Q[F_{\lambda_Q(\mathbf{X})}(u)] = \mathbb{E}_Q[F_{\lambda_Q(\mathbf{Y})}(u)]$, where Q is uniformly picked from $\{1, n\}$. In other words, if the distribution of an eigenvalue picked randomly and uniformly from both matrices is identical, then the two random matrices are regarded to be the same with respect to the MIMO ergodic capacity order.

Although the proposed definition of the MIMO ergodic capacity order is one of many different possible partial orders on matrices, we assert that it is a natural generalization of the ergodic capacity order defined in Section 5.2. This is also elucidated by the fact that the properties M1-M3 and M5 are indeed straight-forward matrix generalizations of properties S1-S3 and S5 respectively. Further, the MIMO ergodic capacity order bears the following connection with the ergodic capacity order defined for random variables:

Proposition : Let $\lambda_Q(\mathbf{X}) \leq_{\mathrm{c}} \lambda_Q(\mathbf{Y})$, where $\lambda_Q(\mathbf{X})$ is an eigenvalue of $\mathbf{X}$ picked uniformly from the set of eigenvalues of $\mathbf{X}$. Then $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y}$. Conversely, if $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y}$, then $\lambda_Q(\mathbf{X}) \leq_{\mathrm{c}} \lambda_Q(\mathbf{Y})$.

Given two MIMO fading systems $\mathbf{X}$ and $\mathbf{Y}$, Proposition 5.4.1 implies that $\mathbf{Y}$ dominates $\mathbf{X}$ in the MIMO ergodic capacity order, if and only if a uniformly randomly selected eigen-channel of $\mathbf{Y}$ has a larger ergodic capacity than that of a uniformly randomly selected eigen-channel of $\mathbf{X}$.

5.4.2 Application

An illustrative example to elucidate the efficacy of the MIMO ergodic capacity order is the M user Gaussian MIMO-MAC, where user i possesses N_t antennas. We assume that only the receiver has CSI, and that each antenna of each user transmits independent signals. Further, each user is allocated the same transmit power ρ per transmit antenna. In this case, the ergodic capacity region $\mathcal{C}_{\text{MM}}(\rho)$ is given by [55]:

$$\mathcal{C}^{X}_{\text{MM}}(\rho) := \Bigg\{ (R_1, \ldots, R_M) : \sum_{i \in \mathcal{S}} R_i \leq \mathbb{E}\left[\log\det\left(\mathbf{I} + \rho g_{\text{MM},\mathcal{S}}(\mathbf{X}_{1:M})\right)\right], \\ \forall \mathcal{S} \subseteq \{1, \ldots, M\} \Bigg\} , \tag{5.19}$$

where $g_{\text{MM},\mathcal{S}}(\mathbf{X}_{1:M}) := \sum_{i \in \mathcal{S}} \mathbf{X}_i$, with $\mathcal{S} \subseteq \{1, \ldots, M\}$. Clearly, when $\mathbf{X}_i$ is assumed to be the variable while viewing all other arguments of $g_{\text{MM},\mathcal{S}}(\cdot)$ as constant matrices, it can be seen that $g_{\text{MM},\mathcal{S}}(\cdot)$ is a Thorin-Bernstein matrix function of $\mathbf{X}_i$, for $i = 1, \ldots, M$. Therefore, through property M4, $g_{\text{MM},\mathcal{S}}(\mathbf{X}_{1:M}) \preceq_{\text{c}} g_{\text{MM},\mathcal{S}}(\mathbf{Y}_{1:M})$, whenever $\mathbf{X}_i \preceq_{\text{c}} \mathbf{Y}_i, i = 1, \ldots, M$. Consequentially, $\mathcal{C}^{X}_{\text{MM}}(\rho) \subseteq \mathcal{C}^{Y}_{\text{MM}}(\rho)$, for $\rho \geq 0$.

5.5 Conclusion

The ergodic capacity order and its properties can be exploited to obtain comparisons of ergodic capacities of composite systems across two different fading channels whose instantaneous SNRs satisfy the ergodic capacity order. For systems such as MRC and EGC which involve multiple instantaneous SNR RVs, we conclude that combining a better set of channels (in the ergodic capacity order) produces a sys-

tem with a higher ergodic capacity. This conclusion is true for all systems whose end-to-end instantaneous SNR belongs to the $\mathcal{CTBF}_m$ set. For systems whose end-to-end SNR does not belong to $\mathcal{CTBF}_m$, component-wise ergodic capacity ordering of instantaneous SNR need not produce a system with a higher ergodic capacity. An example to illustrate this point is the MH-AF relay for which the instantaneous SINR is Pareto-type distributed. An extension of the ergodic capacity order to MIMO systems is also proposed herein. The properties of the ergodic capacity order can be used to compare the capacity regions of systems such as the multi-user MAC in two different fading environments, for both the single and multiple antenna case.

5.6 Proofs: Properties of MIMO Ergodic Capacity Order

We now discuss the proofs of the properties of the MIMO ergodic capacity order. The proofs of the properties S1-S6 of the ergodic capacity order (for scalar RVs) are special cases of Properties M1-M6 respectively, and can be obtained by setting $n = 1$.

Proof of Property M1

Assume $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y}$. Using the identity $\det(\mathbf{I} + \rho\mathbf{X}) = \prod_{i=1}^{n}(1 + \rho\lambda_i(\mathbf{X}))$, we can write

$$\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y} \iff \mathbb{E}\left[\sum_{i=1}^{n}\log(1+\rho\lambda_i(\mathbf{X}))\right] \leq \mathbb{E}\left[\sum_{i=1}^{n}\log(1+\rho\lambda_i(\mathbf{Y}))\right], \forall \rho > 0. \quad (5.20)$$

Multiplying (5.20) by ρ^{-1}, and taking the limit as $\rho \to 0$, it is seen that

$$\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y} \implies \mathbb{E}[\mathbf{tr}\ \mathbf{X}] \leq \mathbb{E}[\mathbf{tr}\ \mathbf{Y}], \quad (5.21)$$

provided the Shannon transforms of $\mathbf{X}$ and $\mathbf{Y}$ exist, and $\mathbb{E}[\mathbf{tr}\ \mathbf{X}] < \infty$ and $\mathbb{E}[\mathbf{tr}\ \mathbf{Y}] < \infty$.

It now follows from (5.20) and (5.21) that

$$\begin{aligned}\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y} \iff & \mathbb{E}\left[\sum_{i=1}^{n} \log(1+t\rho\lambda_i(\mathbf{X}))\mu(t) + a + b\lambda_i(\mathbf{X})\right] \\ & \leq \mathbb{E}\left[\sum_{i=1}^{n} \log(1+t\rho\lambda_i(\mathbf{Y}))\mu(t) + a + b\lambda_i(\mathbf{Y})\right], \\ & \forall a, b \geq 0, \mu(t) \geq 0, \rho > 0, t > 0 \,. \end{aligned} \tag{5.22}$$

Integrating the right hand side of (5.22) over t in the interval $[0, \infty)$ preserves the inequality in (5.22). Therefore,

$$\begin{aligned}\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y} \implies & \mathbb{E}\left[\sum_{i=1}^{n} a + b\lambda_i(\mathbf{X}) + \int_0^\infty \log(1+t\rho\lambda_i(\mathbf{X}))\mu(t)\mathtt{d}t\right] \\ & \leq \mathbb{E}\left[\sum_{i=1}^{n} a + b\lambda_i(\mathbf{Y}) + \int_0^\infty \log(1+t\rho\lambda_i(\mathbf{Y}))\mu(t)\mathtt{d}t\right], \forall \rho > 0 \,. \end{aligned} \tag{5.23}$$

The summand in (5.23) is an arbitrary Thorin-Bernstein function, since a, b, μ are arbitrary and nonnegative. Denoting this Thorin-Bernstein function by g, the direct part of the property is proved by observing from Section 2.1.5 that $\mathbb{E}\left[\sum_{i=1}^{n} g(\lambda_i(\mathbf{X}))\right] = \mathbb{E}\left[\mathbf{tr}\; g(\mathbf{X})\right]$. To prove the converse, choose $g(\mathbf{A}) = \log(\mathbf{I} + \rho\mathbf{A})$.

Proof of Property M2

Let $\mathbf{X}, \mathbf{Y} \in \mathbb{S}_+^n$, and $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y}$. Let $\phi : \mathbb{R} \to \mathbb{R}$ belong to $\mathcal{TBF}$, and $g : \mathbb{R} \to \mathbb{R}$ belong to $\mathcal{CTBF}$. Using the definition of matrix functions, it is easy to see that $f(\mathbf{X}) := \phi(g(\mathbf{X})) \in \mathcal{TBF}$. From Property M1, it is seen that $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y} \iff \mathbb{E}\left[\mathbf{tr}\; \phi(g(\mathbf{X}))\right] \leq \mathbb{E}\left[\mathbf{tr}\; \phi(g(\mathbf{Y}))\right]$. In other words, $g(\mathbf{X}) \preceq_{\mathrm{c}} g(\mathbf{Y})$, which proves the direct part of the property. To see the converse, choose f as the identity map.

Proof of Property M3

Let $\mathbf{X}, \mathbf{Y} \in \mathbb{S}_+^n$, and $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y}$. Using Frullani's formula (2.5), it is evident that an equivalent condition to $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y}$ is given by

$$\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y} \iff \mathbb{E}\left[\int_0^\infty \frac{e^{-s}}{s} \sum_{i=1}^n \left(1 - e^{-\rho s \lambda_i(\mathbf{X})}\right) \mathtt{d}s\right] \leq \mathbb{E}\left[\int_0^\infty \frac{e^{-s}}{s} \sum_{i=1}^n \left(1 - e^{-\rho s \lambda_i(\mathbf{Y})}\right) \mathtt{d}s\right]. \tag{5.24}$$

Commuting the expectation and integral in (5.24), we get

$$\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y} \iff \int_0^\infty \frac{e^{-s}}{s} \mathbb{E}\left[\sum_{i=1}^n e^{-\rho s \lambda_i(\mathbf{X})}\right] \mathtt{d}s \geq \int_0^\infty \frac{e^{-s}}{s} \mathbb{E}\left[\sum_{i=1}^n e^{-\rho s \lambda_i(\mathbf{Y})}\right] \mathtt{d}s. \tag{5.25}$$

Therefore, if $\mathbb{E}\left[\sum_{i=1}^n \exp(-\rho \lambda_i(\mathbf{X}))\right] \geq \mathbb{E}\left[\sum_{i=1}^n \exp(-\rho \lambda_i(\mathbf{X}))\right], \rho > 0$, then $\mathbf{X} \preceq_{\mathrm{c}} \mathbf{Y}$. Finally, observing that $\mathbb{E}\left[\sum_{i=1}^n \exp(-\rho \lambda_i(\mathbf{X}))\right] = \mathbb{E}\left[\mathbf{tr}\, \exp(-\rho \mathbf{X})\right]$, the proof is concluded.

Proof of Property M4

This property is proved using mathematical induction. To begin with, choose a matrix function $\phi \in \mathcal{TBF}$, and $\mathbf{X}_{1:m} := [\mathbf{X}_1, \ldots, \mathbf{X}_m]$ have independent and nonnegative random matrices as components. Assume likewise for $\mathbf{Y}_{1:m} := [\mathbf{Y}_1, \ldots, \mathbf{Y}_m]$. Now, for $m = 1$, Property M4 is true due to Property M2. Next, let us assume Property M4 to be true for sequences of length $m - 1$. Thus, for $g \in \mathcal{CTBF}_m$ we have $g([\mathbf{C}\ \mathbf{X}_{1:m-1}]) \preceq_{\mathrm{c}} g([\mathbf{C}\ \mathbf{Y}_{1:m-1}])$, where $g([\mathbf{C}\ \mathbf{X}_{1:m-1}]) := g(\mathbf{C}, \mathbf{X}_1, \ldots, \mathbf{X}_{m-1})$, and $\mathbf{C} \in \mathbb{S}_+^n$. This implies

$$\mathbb{E}\left[\mathbf{tr}\, \phi\left(g([\mathbf{C}\ \mathbf{X}_{1:m-1}])\right)\right] \leq \mathbb{E}\left[\mathbf{tr}\, \phi\left(g([\mathbf{C}\ \mathbf{Y}_{1:m-1}])\right)\right], \tag{5.26}$$

where we have used Lemma 2.1.1 and Lemma 2.1.2. Next, for sequences of length m, consider

$$\mathbb{E}\left[\mathbf{tr}\,\phi\left(g(\mathbf{X}_{1:m})\right)|\mathbf{X}_1=\mathbf{C}\right]=\mathbb{E}\left[\mathbf{tr}\,\phi\left(g([\mathbf{C}\;\mathbf{X}_{2:m}])\right)\right] \tag{5.27}$$

$$\leq \mathbb{E}\left[\mathbf{tr}\,\phi\left(g([\mathbf{C}\;\mathbf{Y}_{2:m}])\right)\right]=\mathbb{E}\left[\mathbf{tr}\,\phi\left(g(\mathbf{Y}_{1:m})\right)|\mathbf{Y}_1=\mathbf{C}\right]\;, \tag{5.28}$$

where (5.28) follows from (5.27) due to (5.26). Now, taking the expectation with respect to $\mathbf{X}_1$ on the left hand side of (5.27) and the right hand side of (5.28), we get $\mathbb{E}\left[\mathbf{tr}\,\phi\left(g(\mathbf{X}_1,\ldots,\mathbf{X}_m)\right)\right] \leq \mathbb{E}\left[\mathbf{tr}\,\phi\left(g(\mathbf{Y}_1,\ldots,\mathbf{Y}_m)\right)\right]$. Since in the above argument, $\mathbf{X}_1$ is an indeterminate parameter, the same line of reasoning applies when conditioning on any other parameter, and the proof of the property thus follows.

Proof of Property M6

Let $\mathbf{X},\mathbf{Y}\in\mathbb{S}_+^n$, and $\mathbb{E}\left[\log\det\left(\mathbf{I}+\rho\mathbf{X}\right)\right]=\mathbb{E}\left[\log\det\left(\mathbf{I}+\rho\mathbf{Y}\right)\right]$. Using the representation of the log-determinant in terms of the eigenvalues, and (2.12), it is seen that

$$\mathbf{X}=_c\mathbf{Y}\iff\int_0^\infty\frac{\sum_{i=1}^n 1-F_{\lambda_i(\mathbf{X})}\left(u\right)}{1/\rho+u}\mathtt{d}u=\int_0^\infty\frac{\sum_{i=1}^n 1-F_{\lambda_i(\mathbf{Y})}\left(u\right)}{1/\rho+u}\mathtt{d}u\;. \tag{5.29}$$

To see the direct part of the Property, recall the Stieltjes transform of a function of bounded variation is in a one-to-one correspondence with the function, and $\sum_{i=1}^n 1-F_{\lambda_i(\mathbf{X})}\left(u\right)$ is of bounded variation. It is therefore immediate that if $\mathbb{E}\left[\log\det\left(\mathbf{I}+\rho\mathbf{X}\right)\right]=\mathbb{E}\left[\log\det\left(\mathbf{I}+\rho\mathbf{Y}\right)\right]$, then $\sum_{i=1}^n F_{\lambda_i(\mathbf{X})}\left(u\right)=\sum_{i=1}^n F_{\lambda_i(\mathbf{Y})}\left(u\right)$, a.e. To see the converse, let $\sum_{i=1}^n F_{\lambda_i(\mathbf{X})}\left(u\right)=\sum_{i=1}^n F_{\lambda_i(\mathbf{Y})}\left(u\right)$ a.e. Then according to (5.29), $\mathbb{E}\left[\log\det\left(\mathbf{I}+\rho\mathbf{X}\right)\right]=\mathbb{E}\left[\log\det\left(\mathbf{I}+\rho\mathbf{Y}\right)\right]$.

5.7 Motivation and Literature Survey

It is of interest to see if the typical distributions used for multipath, shadowing and composite multipath/shadowing may be unified under a common class with desirable analytical properties. There have been multiple efforts toward unifying fading distributions in the past. One of the unified models - the spherically invariant random process has been proposed as a general model for the fading envelope RV as variance mixture of a Rayleigh RV [56]; a unification proposed in [57] models the instantaneous channel power RV as the product of a gamma RV and a generalized-gamma RV. While this model encompasses many known fading distributions, it does not include some distributions such as the Pareto distribution, which has been used in [31] to account for interference. Another unification in the literature is the generalized gamma distribution (also known as the Stacy distribution [58]), which is shown to correspond to the distribution of the envelope, when the received signal is composed of clusters of multipath components, and the receiver possesses power-law nonlinearity [59]. While this is a physically motivated fading model, it does not include fading distributions such as the Beta-prime distribution, which is used to model the signal to interference ratio in cognitive radio applications [60]. There are several other references where a unified fading model is proposed [61, 62]. However, they do not include some popularly used fading distributions such as the Rician distribution.

In this chapter, we propose to unify fading distributions by modeling their instantaneous channel power as an infinitely divisible (ID) RV. A RV is said to be infinitely divisible, if it can be written as a sum of $n \geq 1$ independent and identically distributed (i.i.d.) RVs, for each n. Intuitively, this means that the channel can be viewed as a (time/frequency/antenna) diversity combining system with n i.i.d. channels, for every $n \geq 1$. Infinitely divisible RVs have many interesting mathematical properties

[63], which are relevant in the performance analysis of wireless systems. Furthermore, almost every distribution used to model multipath, shadowing and composite multipath/shadowing is seen to be included in the class of ID RVs, as shown in this work. Furthermore, most of these distributions are members of a special subclass of ID RVs known as generalized gamma convolutions (GGCs). Consequently, special attention is devoted to this subclass of ID RVs. Formally, GGCs are defined as the weak limits of finite convolutions of Gamma distributions, and have been thoroughly studied in [64, 65, 66, 67]. There are numerous applications of GGCs in financial economics [68, 69, 70] and risk analysis [71, 72]. GGCs possess remarkable analytical structure and closure properties. For example, a sum of independent GGC RVs is a GGC; the product of independent RVs belonging to a subclass of GGC RVs is a GGC. These properties make GGCs an attractive model for multipath fading, shadowing or composite models that incorporate both. Despite possessing properties which are useful in fading system analysis, to the best of our knowledge, no applications of GGCs has been found in the wireless communications literature.

The class of GGCs includes a surprisingly large number of fading models such as the lognormal, Rayleigh, Nakagami-m, generalized gamma (Stacy), Pareto, beta, inverse Gaussian and the positive α-stable distributions. Moreover, composite multipath/shadowing models such as Rayleigh-lognormal, Nakagami-lognormal, and special cases of the spherically invariant random process, and the generalized model proposed in [57] are also GGCs. From a wireless perspective, if a system has GGC distributed instantaneous channel power, then the fading channel is equivalent to a coherent linear combination of (possibly infinitely) many independent branches where the instantaneous fading power is gamma distributed. A subclass of GGCs can also be thought of as a mixture Nakagami-m distribution, where m is random. Unifying instantaneous channel power distributions as GGCs, it is possible to obtain novel gen-

eralized expressions for performance metrics of wireless systems, such as the ergodic capacity and the average SER. Furthermore, this unification reveals new stochastic ordering relations applicable to members of GGC. Lastly, the conditions under which systems composed of multiple fading RVs, such as diversity combining schemes have GGC end-to-end instantaneous channel power, are obtained. This facilitates one to obtain bounds on the performance of such systems.

The rest of the paper is organized as follows. The system model and the performance metrics under consideration are given in Section 5.8. A short exposition of infinite divisibility highlighting connections in wireless communications is presented in Section 5.9. Some examples of fading distributions which are ID are also provided in this section. Performance analysis of GGC fading channels is delegated to Section 5.11. Section 5.12 describes the stochastic ordering of GGC RVs. In Section 5.13, multi-antenna systems such as MRC, EGC and SC are considered.

5.8 System Model

The system model under consideration is the following:

$$y = \sqrt{\rho X}s + w \ , \tag{5.30}$$

where y is the received signal, $\rho \geq 0$ represents the average signal to noise ratio (SNR), s is the transmitted symbol chosen from either one of DPSK or MPSK constellations, w is circularly symmetric additive white Gaussian noise with zero mean and unit variance. In (5.30), X represents the magnitude square of the complex baseband equivalent fading coefficient, and is called as the instantaneous channel power. We normalize $\mathbb{E}[X] = 1$. Unifying the various distributions of X under the assumptions of multipath fading, shadow fading and composite multipath/shadowing is the focus of this work. The unification permits generalized performance analysis of the system

with respect to three metrics: the average SER, the ergodic capacity and the outage probability, which are defined next.

5.8.1 Average Symbol Error Rate

The average SER $\overline{P}_{\mathrm{e}}(\rho)$ is defined as follows:

$$\overline{P}_{\mathrm{e}}(\rho) = \mathbb{E}\left[P_{\mathrm{e}}(\rho X)\right] , \tag{5.31}$$

where $P_{\mathrm{e}}(\rho X)$ is the instantaneous SER, and is dependent on the constellation of choice. For the case of DPSK or MPSK constellation, the instantaneous SER is c.m., and can be written as [50]:

$$P_{\mathrm{e}}(\rho x) = \int_0^{\infty} \exp(-\rho u x)\eta(u)\mathtt{d}u , \tag{5.32}$$

where $\eta(\cdot)$ is given by

$$\eta(u) = \begin{cases} \frac{1}{2} I\left(u = 1\right), & \text{DPSK} \\ \frac{\sqrt{\beta_{\mathrm{PSK}}}}{2\pi} \frac{I\left(u \geq \frac{\beta_{\mathrm{PSK}}}{\sin^2\left((M-1)\frac{\pi}{M}\right)}\right)}{u\sqrt{u-\beta_{\mathrm{PSK}}}}, & \text{MPSK} \end{cases} , \tag{5.33}$$

where $\beta_{\mathrm{PSK}} = \sin^2(\pi/M)$. This representation of the SER will be used in Section 5.12.

The analysis of the asymptotic SER is considered in Section 5.11.1. In the asymptotically high SNR regime, $\overline{P}_{\mathrm{e}}(\rho)$ is given by

$$\overline{P}_{\mathrm{e}}(\rho) \sim c_Q l(\rho^{-1})\rho^{-D} , \rho \to \infty , \tag{5.34}$$

where $l(x)$ is slowly varying at 0. In (5.34), the diversity order D is the asymptotic slope of the average SER on a log-log plot versus the average SNR, and is defined as [73]

$$D = -\frac{1}{\log t} \log \lim_{\rho\to\infty} \frac{\overline{P}_{\mathrm{e}}(\rho t)}{\overline{P}_{\mathrm{e}}(\rho)}, \; t > 0 , \tag{5.35}$$

which is also the variation exponent of (5.31) at ∞. The term c_Q in (5.34) is a modulation dependent constant, and is given by $c_Q = \pi/4$ for the case of DPSK [73], and

$$c_Q = (\pi \sin^{2D}(M^{-1}\pi))^{-1} \int_0^{(1-M^{-1})\pi} \sin^{2D}\theta \mathtt{d}\theta \ , \tag{5.36}$$

for the case of MPSK [73]. We also define an asymptotic quantity known as the array gain, which represents the shift of the average SER curve to the right on a log-log plot, and is defined as

$$G = c_Q \lim_{\rho\to\infty} \rho^D \overline{P}_{\mathrm{e}}\ (\rho)\ . \tag{5.37}$$

Although the limit in (5.37) may not exist every fading distribution, it will be shown in Section 5.11.1 that the limit exists for GGCs satisfying certain conditions, and is in fact, a finite constant.

5.8.2 Outage Probability

The outage probability for the system (5.30) is defined as the probability that the instantaneous SNR is less than a fixed threshold value, and is given by

$$P_{\mathrm{out}}(x/\rho) = \Pr[X \leq x/\rho] \ , \tag{5.38}$$

where x is a constant.

5.9 Infinitely Divisible Fading Distributions

In what follows, a short exposition of infinite divisibility and some of its subclasses is provided, with some examples of popularly used fading distributions which are ID.

5.9.1 Infinite Divisibility

A probability distribution is said to be infinitely divisible (ID) if, for each $n \geq 1$, it can be decomposed into n identical convolution factors. Some examples of

such distributions are the Gamma, the Poisson and the Gaussian distributions. In this paper, we refer to RVs with ID distributions as ID RVs. Non-negative ID RVs are characterized by the fact that the Laplace exponents of such RVs are Bernstein functions. The set of all nonnegative ID RVs is denoted by $\mathcal{I}$. Nonnegative ID RVs are very relevant from a wireless communications context, because it is intuitive to view fading as the resulting effect of coherent combination of contributions from multiple different sources. Indeed, almost all known instantaneous channel power distributions are ID, with examples listed in Section 5.9.4. The infinite divisibility of the lognormal distribution, proved by Thorin [74] revealed a very rich class of ID probability distributions known as generalized gamma convolutions (GGC), which possess extremely useful mathematical properties applicable to wireless systems. In addition, several popularly used fading distributions are not only ID, but also GGC. This fact motivates us to consider this class of ID distributions, which is discussed next.

5.9.2 Generalized Gamma Convolutions

The subset of non-negative ID distributions for which the Laplace exponent belongs to $\mathcal{TBF}$, constitutes the set $\mathcal{E}$ of GGC probability distributions. In other words, the LT of the RV admits the form

$$\phi(s) = \exp\left(-as - \int_0^\infty \log(1 + s/u)\mu(\mathtt{d}u)\right) , \tag{5.39}$$

where the parameters $a \geq 0$, and $\mu(\cdot)$ (nonnegative measure on $[0, \infty)$) uniquely characterize the GGC distribution. In this paper, we call the Thorin measure of the Laplace exponent of the GGC (e.g. μ in (5.39)) as the Thorin measure of the GGC.

In a wireless context, the Thorin measure of a GGC has a physical interpretation: Consider the case where a continuum of fading channels with Gamma distributed

instantaneous power is coherently summed up. If the sequence of shape parameters corresponding to each rate parameter u in $[0, \infty)$ of the Gamma distributed channels is $\mu(u)$, then the Thorin measure of the GGC channel is μ.

In what follows, we will restrict our attention to nondegenerate GGCs with $a = 0$, because a merely represents a translation of the instantaneous channel power PDF, and performance analysis for such channels can be obtained from systems where $a = 0$ using straightforward algebra.

The variation exponent of the LT of the GGC near infinity will correspond to the diversity order, and this quantity relates to the Thorin measure as follows:

Lemma 5.9.1. *[64, Theorem 4.1.4] The LT of a GGC is regularly varying at ∞, and the variation exponent equals the Thorin mass. That is,* $\lim_{\rho\to\infty} \phi(t\rho)/\phi(\rho) = t^{-c}$, *where*

$$c = \int_0^\infty \mu(\mathtt{d}s) \tag{5.40}$$

The variation exponent of the LT of a GGC near ∞ will have applications in asymptotic performance analysis, as discussed in Section 5.11.

The Thorin measure also determines the expected value of a GGC RV:

Proposition : Let X be a GGC. Then $\mathbb{E}[X] = \int_0^\infty \mu(\mathtt{d}u)/u$, if the integral is finite.

Proof. It is easy to show for any nonnegative RV that $\mathbb{E}[X] = -\lim_{\rho\to 0} \mathtt{d}\phi_X(\rho)/\mathtt{d}\rho$. Since the LT of a GGC is of the form $\phi_X(\rho) = \exp(-\gamma(\rho))$, $\mathbb{E}[X] = \lim_{\rho\to 0} \mathtt{d}\gamma(\rho)/\mathtt{d}\rho$, where $\gamma(\cdot)$ is the Laplace exponent of the GGC. Representing γ in terms of the Thorin measure,

$$\mathbb{E}[X] = \lim_{\rho\to 0} \int_0^\infty \frac{1/u}{1+\rho/u} \mu(\mathtt{d}u) \,, \tag{5.41}$$

where μ is the Thorin measure. The limit and integral can be interchanged whenever

$$\int_0^\infty u^{-1}\mu(\mathtt{d}u) < \infty \ , \tag{5.42}$$

because under this condition dominated convergence is satisfied. Assuming (5.42) to be true, the limit and integration are interchanged, to yield the Proposition. □

If X represents the instantaneous channel power in a fading system, then Proposition 5.9.2 gives the average channel power of the system. Hereafter, we assume that the mean of the GGC is normalized to 1, as per the discussion after (5.30). In doing so, we ensure that ρ in (5.30) reflects the signal to noise ratio of the system.

Next, an alternate characterization of GGCs is in terms of gamma distributions is noted. Recall that the gamma density with shape parameter α and rate parameter θ (represented as $Gamma(\alpha, \theta)$) is given by

$$f_X(x) = \frac{x^{\alpha-1}\theta^\alpha}{\Gamma(\alpha)} \exp(-\theta x) \ , \tag{5.43}$$

where $\Gamma(\cdot)$ is the Gamma function. Then, an arbitrary GGC is a weak limit of a sequence of convolutions of gamma distributions with suitably chosen parameters. It is straightforward to see that, if X_1 and X_2 are two GGCs with Thorin measures μ_1 and μ_2, then $X_1 + X_2$ is a GGC with a Thorin measure $\mu_1 + \mu_2$.

Some useful properties of non-degenerate GGCs are now presented.

The density of a GGC always exists [64, Theorem 4.1.2, Theorem 4.1.3], and can be written in terms of c.m. functions as follows:

Lemma 5.9.2. *[64, Theorem 4.1.1, Theorem 4.1.2] The density of a GGC with Thorin mass $\alpha < \infty$ can be written in the form*

$$f_X(x) = x^{\alpha-1}h(x) \ , \tag{5.44}$$

where $h(\cdot)$ is a c.m. function, which is slowly varying at 0. Moreover, $h(0+)$ is finite if and only if $\int_1^\infty \log u \mu(\mathtt{d}u) < \infty$, and in this case

$$h(0+) = \frac{1}{\Gamma(\alpha)} \exp\left(\int_0^\infty \log u \mu(\mathtt{d}u) \right) . \tag{5.45}$$

The representation of the PDF of a GGC in the form (5.44) is particularly useful to obtain asymptotic performance metrics as shown in Section 5.11.1. GGCs are also mixtures of Gamma distributions, as summarized in the following Lemma:

Lemma 5.9.3. *[64, Theorem 4.1.1] If X is a GGC with Thorin mass $\alpha < \infty$, then $X = AZ$, where $Z \sim Gamma(\alpha, 1)$ and $A \geq 0$ are independent. Moreover, $h(x)$ in (5.44) is given by*

$$h(x) = (\Gamma(\alpha))^{-1} \mathbb{E}\left[A^{-\alpha} \exp(-x/A) \right] . \tag{5.46}$$

Interpreting X as the instantaneous channel power of a GGC channel, from Lemma 5.9.3 it is straightforward to see that every GGC channel with finite Thorin mass is equivalent to a variance mixture of a Nakagami-m channel. This is because the instantaneous channel power corresponding to Nakagami-m fading is Gamma distributed [34]. It is shown in Section 5.9.4 that several fading distributions such as Nakagami-m, Pareto and generalized gamma distributions have instantaneous channel powers which are GGC RVs, and the corresponding Thorin masses are finite.

5.9.3 Hyperbolically Completely Monotone RVs

A subset of $\mathcal{E}$, consisting of the distributions for which the density is a hyperbolically completely monotone (HCM) function is denoted by $\mathcal{H}$, and the RVs corresponding to distributions in this set are called HCM RVs. A function g is defined to

be a HCM function if $g(uv)g(u/v)$ can be written as [64, p. 68]

$$g(uv)g(u/v) = \int_0^\infty \exp(-\lambda u(v+v^{-1}))K(\mathtt{d}\lambda; u)\ , K(\mathtt{d}\lambda; u) \geq 0\ , \tag{5.47}$$

for each $u > 0$. HCM RVs enjoy the following properties under products and quotients of RVs, which will be useful in obtaining closure results for wireless systems involving multiple fading RVs:

Lemma 5.9.4. *If X_1, X_2 are independent HCM RVs, and Y is a GGC independent of X_1 and X_2, then*

(i) $X_1^q \in \mathcal{H}$, *for* $|q| \geq 1$.

(ii) $X_1 X_2 \in \mathcal{H}$.

(iii) $X_1 Y \in \mathcal{E}$.

The density of a HCM RV admits the following canonical characterization

Lemma 5.9.5. *[64, p. 81] The density function of a HCM RV admits the canonical form*

$$f_X(x) = Cx^{\alpha-1}p_1(x)p_2(1/x) \tag{5.48}$$

where $\alpha \in \mathbb{R}$, and

$$p_j(x) = \exp\left(-b_j x - \int_0^\infty \log\left(\frac{x+y}{1+y}\right) T_j(\mathtt{d}y)\right)\ , \tag{5.49}$$

with $b_j \geq 0$, and $T_j(\mathtt{d}y)$ being non-negative measures on $(0,\infty)$ satisfying $\int_0^\infty (1+y)^{-1} T_j(\mathtt{d}y) < \infty$, $j = 1, 2$.

In the case of HCM RVs, the Thorin mass can be represented in terms of the parameters α, b_j and T_j as

Lemma 5.9.6. *[64, p. 74] If X is a HCM RV with canonical PDF parameters α, b_j and T_j, $j = 1, 2$, then*

$$\int_0^\infty \mu(\mathtt{d}s) = \begin{cases} \alpha + T_2([1,\infty]) & \text{, if } b_2 = 0 \\ \infty & \text{, if } b_2 \neq 0 \end{cases}. \tag{5.50}$$

Many RVs relevant to wireless communications are HCM RVs. For example, the Gamma RV, the positive stable RV and the lognormal RV are HCM RVs.

In this work, the conditions under which the Thorin mass of a sum or product of GGC's is finite is required. To the best of our knowledge, these conditions were not found elsewhere in the literature, although it is not too difficult to derive them.

Proposition : Let X_1, X_2, X_3 be independent nonnegative RVs, where X_1 and X_2 are GGC's and X_3 is a HCM RV. Then

(i) The Thorin mass of $X_1 + X_2$ is finite if and only if the Thorin masses of both X_1 and X_2 are finite.

(ii) The Thorin mass of $X_1 X_3$ is finite if the Thorin mass of either X_1 or X_3 is finite.

(iii) The Thorin mass of X_3^{-1} is finite if the canonical parameters of X_3 in (5.48) satisfy $\alpha < \infty$, $b_1 = 0$ and $T_1([1,\infty)) < \infty$.

5.9.4 Examples

In this work, it is proposed to view instantaneous channel power distributions as members of $\mathcal{I}$. To justify that the $\mathcal{I}$ class is indeed a unification of fading distributions, many of the popularly used fading channels are shown to have ID instantaneous channel powers. Among these channels, distributions for which the instantaneous channel power is a GGC or HCM RV are identified and summarized in Table 5.1, as this fact will be useful in the analysis of systems, where the overall instantaneous

Name	**ID**	**GGC**	**HCM**
Rician	Yes	No	No
Nakagami-m	Yes	Yes	Yes
Hoyt	Yes	Yes	Yes
Generalized gamma distribution	Yes	Yes	Yes
Lognormal	Yes	Yes	Yes
Generalized-K	Yes	Yes	Yes
$N*$Nakagami	Yes	Yes	Yes
Weibull	Yes	Yes	Yes
Positive stable	Yes	Yes	some cases
Pareto	Yes	Yes	Yes
Rayleigh-lognormal	Yes	Yes	Yes
Nakagami-lognormal	Yes	Yes	Yes
Spherically invariant random process	Yes	some cases	some cases

Table 5.1: Commonly used fading distributions, and whether their instantaneous channel power belongs to $\mathcal{I}$, $\mathcal{E}$ or $\mathcal{H}$.

channel power RV involves a sum or product of two or more RVs (See Section 5.13). Explicit examples of RVs which are ID, GGC and HCM are now listed.

Rician-K Fading

Rician-K distribution is used to model the fading envelope in certain line-of-sight scenarios [2, p. 21]. For this case, if X is a nonnegative RV corresponding to the instantaneous channel power, then $\sqrt{X}$ is Rice distributed with parameter K. The LT of X is given by [2, p. 19]

$$\phi_X(s) = \frac{1+K}{1+K+s} \exp\left(-\frac{Ks}{1+K+s}\right) . \tag{5.51}$$

It is now shown that X is an ID RV. To begin with, the Laplace exponent of X is obtained from (5.51) as

$$\gamma_X(s) = \frac{Ks}{1+K+s} + \log\left(1 + \frac{s}{1+K}\right) . \tag{5.52}$$

Using Frullani's representation for $\log(1+x)$ [19, p. 6], it is straightforward to show that $\gamma_X(s)$ can be written in the form (2.4) with $a = b = 0$ and

$$\tau(s) = \exp(-(1+K)s)\left(K(1+K) + \frac{1}{s}\right) . \tag{5.53}$$

Thus, $\gamma_X(s)$ is a Bernstein function, and consequently, X is an ID RV. However, this is not a GGC because the Laplace exponent $\gamma_X(\cdot)$ is not a Thorin-Bernstein function. To see this, recall from Chapter 2.1.3 that a Bernstein function with Levy measure τ is a Thorin-Bernstein function, if and only if $s\tau(s)$ is a c.m. function. However, by differentiating $s\tau(s)$, where $\tau(s)$ is defined in (5.53), it is seen that the first derivative is not non-positive at all values of s, which shows that $s\tau(s)$ is not c.m..

The instantaneous channel power corresponding to Rician-K fading is therefore an ID RV, but not a GGC.

Spherically Invariant Random Process

The spherically invariant random process (SIRP) model has been proposed as a unified model for the fading envelope distribution in the literature [56]. In this case, the instantaneous channel power X is given by

$$X = AE , \tag{5.54}$$

where E is an exponential RV, and A is a positive valued RV.

It is now shown that X for this case is always an ID RV, and is GGC and HCM for certain special cases. To see that X is ID, observe that X is a variance mixture

of exponentials, which is ID, according to [64, Theorem 2.4.3]. Not every member of SIRP is a GGC. A simple counterexample is the case of Rician fading, which is a member of SIRP according to [56], and not a GGC as seen in Section 5.9.4. The special cases of SIRP which are GGC RVs and HCM RVs are now considered. X is a GGC, whenever A is a GGC, according to Lemma BLAH. Also, X is a HCM RV for all such members of SIRP where A is a HCM RV.

Nakagami-m Fading

Nakagami-m fading is used to model short-term fading effects in line of sight channels. In addition, the Gamma distribution has also been used to model long-term fading effects in the literature [75]. Moreover, the exponential distribution, which is a special case of a Gamma RV with unit shape and rate parameters, arises as the instantaneous fading power distribution in Rayleigh fading channels.

The instantaneous channel power X of Nakagami-m fading is Gamma distributed with shape m and rate m[34]. In this case, X is a HCM, GGC and ID RV. To see that X is a GGC, observe that the LT of X is given by $\phi_{_X}(s) = \exp(-m\log(1+s/m))$, which satisfies the representation of the LT of a GGC specified in (5.39). It is then straightforward to show using (5.39) that the Thorin measure in this case, is given by

$$\mu(u) = mI(u = m) \ , \tag{5.55}$$

and the Thorin mass is m, as can be verified by rewriting the PDF in the form (5.44). Further, $h(\cdot)$ is given by $h(x) = m\exp(-mx)/\Gamma(m)$. The proof of X being a HCM RV is given in [64, p. 75]. The canonical parameters of (5.48) are given by $\alpha = m$, $b_1 = m$, $b_2 = 0$, $T_j = 0, j = 1, 2$. Therefore, X is a HCM, GGC and ID RV

Nakagami-q (Hoyt) Fading

Nakagami-q (Hoyt) fading is observed in scenarios such as satellite communications subject to ionospheric scintillation [76]. Special cases of Hoyt fading are Rayleigh fading ($q = 1$) and one-sided Gaussian ($q = 0$).

The in-phase and quadrature components of the complex baseband fading RV for the case of Nakagami-q fading are independent zero mean Gaussian RVs with unequal variances, such that the ratio of the quadrature variance to the in-phase variance equals q [77, 78]. Further, the average channel power is unity if the sum of the variances is 1. As a result, the instantaneous channel power X can be written as

$$X = \frac{1}{1+q^2} X_1 + \frac{q^2}{1+q^2} X_2 \,, \tag{5.56}$$

where X_1 and X_2 are independent (central) chi-squared RVs with 1 degree of freedom, and $0 \leq q \leq 1$.

It is now argued that X is a GGC and hence also ID. Toward this end, first observe that X_i, $i = 1, 2$ is a GGC, because its Laplace exponent is given by $\gamma(s) = (1/2)\log(1 + 2s)$, which is a Thorin-Bernstein function. It then follows from (5.56) that X is a GGC, because GGCs are closed under nonnegative scaling of RVs and addition of RVs. Moreover, using the identity $\phi_{aX}(s) = \phi_X(as)$ for any nonnegative RV X, it can be shown that the Thorin measure is given by

$$\mu(u) = \frac{1}{2(1+q^2)} I\left(u = \frac{1+q^2}{2}\right) + \frac{q^2}{2(1+q^2)} I\left(u = \frac{1+q^2}{2q^2}\right) \,, \tag{5.57}$$

and the Thorin mass is 1. It is also noted that X is not HCM, because the Thorin measure of a HCM RV is allowed to have at most one atom [64, p. 88], which is not the case for X, as observed from (5.57). Therefore, X for Hoyt fading is a GGC and ID RV, but not a HCM RV.

General Lognormal Distribution

The lognormal distribution is commonly used to model the long-term shadowing effect in wireless channels [2].

The density of the general lognormal RV is given by

$$f_X(x) = cx^{\mu/\sigma^2-1} \exp\left(-\frac{(\log x)^2}{2\sigma^2}\right) , \tag{5.58}$$

with $\mu \in \mathbb{R}, \sigma > 0$, and $c > 0$ is a normalizing constant. It has been shown in [64, p. 74] that the general lognormal distribution belongs to $\mathcal{H}$, and hence also $\mathcal{E}$ and $\mathcal{I}$. Although the Thorin measure is unknown, the canonical parameters of (5.48) are known to be $b_1 = b_2 = 0$, $\alpha = \mu/\sigma^2$ and $T_j(u) = \sigma^{-2}u^{-1}, u > 1$ [64, p. 74], and the Thorin mass can be calculated from (5.50) to be ∞.

Generalized Gamma Distribution (Stacy Distribution)

The instantaneous channel power X is generalized Gamma distributed, when the envelope of the fading amplitude is modelled as a generalized Nakagami-m RV [75]. The Weibull distribution is a special case of the density in (5.60), which has been used to approximate the multipath wireless channel from channel measurements [79]. Other distributions which are special cases of the generalized Gamma distribution are the inverse-Gaussian distribution and the Nakagami-m distribution. Moreover, the generalized gamma distribution arises as the instantaneous channel power distribution in Nakagami-m fading channels with receiver non-linearity, where the non-linearity is captured as a power parameter, since X can be written as [64, p. 13]

$$X = Y^{1/r} , \tag{5.59}$$

where $Y \sim Gamma(\epsilon/r, c_2)$.

The density of X is given by

$$f_X(x) = c_1 x^{\epsilon-1} \exp\left(-c_2 x^r\right) \ , \tag{5.60}$$

where $0 < |r| \leq 1$, $c_1, c_2 \geq 0$, and $\epsilon/r > 0$.

It is now shown that this distribution belongs to $\mathcal{H}$ and consequently $\mathcal{E}$ and $\mathcal{I}$. Recall that the Gamma distribution belongs to $\mathcal{H}$, as shown in Section 5.9.4. Using Lemma 5.9.4, it is then seen that the generalized gamma distribution also belongs to $\mathcal{H}$, and is therefore a GGC. While the Thorin measure for the general case is unknown, the Thorin mass is equal to ϵ if $r > 0$, and ∞ otherwise, as can be verified by representing (5.60) in the form (5.44), and further, the function $h(\cdot)$ is given by $h(x) = c_1 \exp(-c_2 x^r)$.

Product of Generalized Gamma Random Variables

The product of independent generalized Gamma RVs has been proposed as a unified fading model in the literature [62]. As a special case, a product of N independent Gamma RVs is obtained, which is the instantaneous channel power distribution in the generalized-K fading model [80]. Another special case of this distribution is the $N*$Nakagami fading model [61, 81], where the envelope is a product of N independent but not necessarily identically distributed Nakagami-m RVs.

Consider X to be a product of N independent but not necessarily identically distributed generalized Gamma distributed RVs. This distribution is HCM as observed from an application of Lemma 5.9.4 and the fact that the generalized Gamma distribution belongs to $\mathcal{H}$. As a consequence, the distribution is also in $\mathcal{E}$ and $\mathcal{I}$.

Positive Stable Distribution

Positive stable distributions have been used to model the interference at the primary receiver in a cognitive radio network, when the interfering secondary terminals are distributed in a Poisson field, and there is a guard zone around the primary receiver [82].

The positively skewed stable distribution is a heavy tailed distribution, which is characterized by its LT as $\phi_X(s) = \exp(-s^r)$, $0 < r \leq 1$. No closed form expression for the distribution or density is known in general. Nevertheless, it is known that this distribution is a GGC [64, p. 35], with Thorin measure given by

$$\mu(u) = \frac{r \sin r\pi}{\pi} u^{r-1} , \tag{5.61}$$

and Thorin mass equal to ∞. It has been established in the literature that this distribution belongs to $\mathcal{H}$ when $r \in (0, 1/4) \cup [1/3, 1/2]$ [83].

Pareto Distribution

Pareto RVs are heavy-tiled distributions used to model signal to interference ratios in interference dominated scenarios. A Pareto RV can be written in the form [64, p. 14]

$$X = \left(\frac{X_1}{X_2}\right)^{\frac{1}{r}} , \tag{5.62}$$

where $X_j \sim Gamma(k_j, 1)$, $j = 1, 2$. The fact that Pareto RVs are ratios of Gamma RVs raised to a non-negative power, as shown in (5.62) leads to many wireless systems with Pareto instantaneous channel power. For example, consider a system with a transmitter, receiver and an interfering terminal. Suppose the channel between the transmitter and the receiver, and that between the interferer and receiver are both Nakagami-m channels (with possibly different parameters), and the receiver is non-linear with the non-linearity captured in the form of a power parameter. Then from

(5.62), it is clear that the instantaneous channel power is a Pareto RV. It is also noted that special cases of the Pareto distribution with ($k_1 = k_2 = 1, r > 0$) have been used to model the instantaneous signal to interference power in interference dominated networks [31].

The Pareto distribution given by the density

$$f_X(x) = \frac{|r|}{B(k_2, k_1)} \frac{x^{k_1 r - 1}}{(1 + x^r)^{k_1 + k_2}}, \tag{5.63}$$

where $B(a, b) := \int_0^1 t^{a-1}(1-t)^{b-1}\mathtt{d}t$ is the Beta function, $r \in [-1, 1]$, $k_j > 0$, $j = 1, 2$.

The Pareto distribution belongs to $\mathcal{H}$. This is because Gamma RVs are HCM, and using Lemma 5.9.4 with (5.62), it follows that a Pareto RV is HCM. The Thorin measure is unknown in the general case, however the Thorin mass is obtained from Lemma 5.9.6 as $k_1 r$.

As discussed above, almost all of the fading distributions with ID instantaneous channel power also belong to $\mathcal{E}$. It is therefore reasonable to focus on channels with instantaneous channel power in $\mathcal{E}$. Hereafter, we refer to channels with GGC distributed channel power as GGC channels.

5.10 GGC Fading Channels as Gamma Mixtures

Let X be a GGC with finite Thorin mass α. As discussed in Lemma 5.9.3, X can be written as AZ, where Z is Gamma distributed with parameters $(\alpha, 1)$, and A is a nonnegative RV independent of Z. Using the property $tZ \sim Gamma(\alpha, 1/t)$ applicable to Gamma RVs, it is easy to see that X can be written as $\tilde{A}\tilde{Z}$, where $\tilde{Z}$ is Gamma distributed with parameters (α, α), and $\tilde{A}$ is a nonnegative RV independent of $\tilde{Z}$. This representation shows that a GGC with channel with finite diversity order is a Nakagami-m fading channel with shadowing, where $m = \alpha$. This is because the instantaneous channel power of a Nakagami-m channel with $m = \alpha$ has the same

distribution as $\tilde{Z}$. Moreover, the average channel power of the shadowing RV is 1, since it has been assumed that $\mathbb{E}[X] = 1$.

The distribution of the shadowing RV $\tilde{A}$ is not trivial to obtain, when only the distribution of X is known. In this case, we adapt the approach used to obtain the mixing distribution for mixture of exponential distributions [56]. This method is based on the fact that the Mellin transform of a product of independent RVs is the product of the Mellin transforms of the RVs, and that the distribution and its Mellin transform pair are unique. The Mellin transform of a nonnegative RV $\tilde{A}$ is defined as

$$M_{\tilde{A}}(s) = \int_0^\infty f_{\tilde{A}}(x) x^{s-1} \mathtt{d}s \ , \tag{5.64}$$

and $M_{\tilde{A}\tilde{Z}}(s) = M_{\tilde{A}}(s) M_{\tilde{Z}}(s)$. Therefore, the PDF of $\tilde{A}$ can be obtained using the inverse Mellin transform as follows:

$$f_{\tilde{A}}(x) = M^{-1}\left[\frac{M_X(s)}{M_{\tilde{Z}}(s)}\right] \ . \tag{5.65}$$

The Mellin transform of $\tilde{Z}$ can be evaluated as $m^{1-s}\Gamma(s+\alpha-1)/\Gamma(\alpha)$ [84, p. 312]. Thus, if the Mellin transform of the GGC X is known, then the density of X can be obtained in terms of Meijer-G functions [56].

5.11 Performance Metrics for GGC Channels

In this Section, asymptotic expressions for average SER of GGC channels in terms of the Thorin measure and/or the function $h(\cdot)$ defined in (5.44), are obtained.

5.11.1 Asymptotic Symbol Error Rate

The average SER performance of a GGC channel at high SNR is now considered, in order to obtain the diversity order and array gain in terms of the Thorin measure.

Proposition : Let a GGC fading channel have Thorin mass $\alpha < \infty$. Then for DPSK/MPSK, we have $\overline{P}_{\mathrm{e}}(\rho) \sim c_Q h(\rho^{-1})\rho^{-\alpha}$, as $\rho \to \infty$, where c_Q depends on the modulation scheme, and $h(\cdot)$ is as defined in (5.44) or (5.46).

Proof. According to [73, Theorem 4], $\overline{P}_{\mathrm{e}}(\rho)$ can be written as $c_Q l(\rho^{-1})\rho^{-D}$, as $\rho \to \infty$ for DPSK and MPSK, if the asymptotic density of X has the form $f_X(x) = x^{D-1}l(x)$, as $x \to 0$. GGCs with finite Thorin mass satisfy this condition as seen from Lemma 5.9.2. □

Using the definition of diversity order from (5.35), it is observed from Proposition 5.11.1 that the diversity order of a GGC channel is as follows.

Corollary 5.11.1. *The diversity order of DPSK/MPSK over a GGC fading channel with Thorin mass $\alpha < \infty$ is equal to α.*

The asymptotic SER obtained in Proposition 5.11.1 applies to any GGC fading channel with finite diversity order, and shows the dependence of the parameters of the GGC on the asymptotic SER parameters. For GGC channels satisfying a certain condition on the Thorin measure, the asymptotic SER can be written in the form $\overline{P}_{\mathrm{e}}(\rho) \sim G\rho^{-\alpha}$.

Proposition : Let the Thorin measure of a GGC fading channel with finite Thorin mass α satisfy $\int_1^\infty (\log u)\mu(\mathtt{d}u) < \infty$. Then $\overline{P}_{\mathrm{e}}(\rho) \sim (G\rho)^{-\alpha}$, as $\rho \to \infty$, where

$$G = \left[\frac{c_{\mathrm{M}}}{\Gamma(\alpha)} \exp\left(\int_0^\infty (\log u)\mu(\mathtt{d}u)\right)\right]^{-\frac{1}{\alpha}} . \tag{5.66}$$

Proof. From Proposition 5.11.1,

$$\lim_{\rho\to\infty} \frac{c_{\mathrm{M}}^{-1}\rho^{\alpha}\overline{P}_{\mathrm{e}}(\rho)}{h(\rho^{-1})} = 1 . \tag{5.67}$$

The limit of the ratio in (5.67) can be written as a ratio of limits of the numerator and denominator, if and only if $\int_1^\infty (\log u)\mu(\mathtt{d}u) < \infty$, because this is the condition under which $\lim_{\rho\to\infty} h(\rho^{-1}) < \infty$, as can be seen through an application of Lemma 5.9.2. Therefore, (5.67) simplifies to

$$\lim_{\rho\to\infty} c_{\mathrm{M}}^{-1}\rho^{\alpha}\overline{P}_{\mathrm{e}}\ (\rho) = \lim_{\rho\to\infty} h(\rho^{-1})\ . \tag{5.68}$$

The right hand side of (5.68) is given by $\Gamma(\alpha)^{-1}\exp(\int_0^\infty (\log u)\mu(\mathtt{d}u))$, as seen from Lemma 5.9.2. Substituting this in (5.68) and rearranging,

$$\lim_{\rho\to\infty} \frac{\overline{P}_{\mathrm{e}}\ (\rho)}{\rho^{-\alpha}c_{\mathrm{M}}\Gamma(\alpha)^{-1}\exp(\int_0^\infty (\log u)\mu(\mathtt{d}u))} = 1\ . \tag{5.69}$$

This proves the Proposition. □

According to Proposition 5.11.1, the slow varying function $h(x)$ in Proposition 5.11.1 becomes a constant at 0. This conclusion is not necessarily true for fading distributions which are not GGCs, or are GGCs which do not satisfy $\int_1^\infty (\log u)\mu(\mathtt{d}u) < \infty$. An example of a fading distribution for which the slow varying function h does not become a constant at 0 is the case of generalized K-fading, as discussed in [73].

It is apparent from (5.66) that the array gain is inversely proportional to $h(0+)$, where $h(\cdot)$ is as defined in (5.44), because the term $\Gamma(\alpha)^{-1}\exp(\int_0^\infty (\log u)\mu(\mathtt{d}u))$ equals $h(0+)$, according to Lemma 5.9.2. The array gain will be useful in calculating the performance difference between two GGC fading channels at high SNR, as discussed in Section 5.12.3.

5.11.2 Outage Probability

Since the outage probability is equal to the distribution function of the instantaneous channel power, the following discussion is useful for cases when the distribution is not available in closed form. The outage probability of a GGC fading channel can

be shown to be a logconcave function, since the density of a GGC is a logconcave function [64, p. 67]. In what follows, an asymptotic expression for the outage probability demonstrating the dependence on Thorin measure is presented.

Proposition : Consider a GGC channel with Thorin mass $\alpha < \infty$. Then the asymptotic outage probability $P_{\text{out}}(\cdot)$ is given by

$$P_{\text{out}}(x) \sim x^{\alpha} h(x)/\alpha, \tag{5.70}$$

as $x \to 0$, where $h(\cdot)$ is as defined in (5.44) or (5.46).

Differentiating the expression the outage probability in Proposition 5.11.2 with respect to α, it can be seen that the outage probability for asymptotically low threshold values monotonically decreases with the Thorin mass. This can be justified intuitively, since the Thorin mass represents the diversity order of the fading channel. A larger value of diversity order corresponds to an increased number of independent fading paths for signal reception, which reduces the probability of outage.

5.12 Stochastic Ordering of GGC Distributions

In this Section, the stochastic ordering of GGC fading channels is considered. This will help in comparing two GGC fading channels based on general performance metrics.

5.12.1 Laplace Transform Ordering of GGC

The comparison of GGC fading channels based on metrics which are either completely monotone (such as SERs), or possess a completely monotone derivative (such as the ergodic capacity) is now considered. Toward this end, the Laplace transform ordering framework proposed in [39] is employed. LT ordering between a pair of

instantaneous channel power distributions implies that the average SER of a constellation with c.m. SER, such as the case with MPSK and MQAM, will be ordered at all values of average SNR.

In the context of LT ordering of GGCs, a duality between the Shannon transform ordering of the Thorin measures with the LT ordering of the fading distributions is straightforward to see:

Proposition : Let X and Y be two GGCs with Thorin measures μ_X and μ_Y. We have $X \leq_{\mathrm{Lt}} Y \iff \mu_X \leq_{\mathrm{S}} \mu_Y$.

Proof. The proof of this proposition follows from the characterization of the LT of a GGC. □

The connection between the LT order and the Shannon transform order as suggested by Proposition 5.12.1 can be exploited in obtaining new ordering relations between GGC fading distributions, by using the properties of one stochastic order to benefit the other. For instance, by observing that $g(x) = \log(1 + \rho x)$ is a Bernstein function for $\rho \geq 0$, it is seen that $\mu_X \leq_{\mathrm{Lt}} \mu_Y \implies \mu_X \leq_{\mathrm{S}} \mu_Y$, for cases when the Thorin masses of X and Y are identical. Now, using Proposition 5.12.1, it is concluded that $\mu_X \leq_{\mathrm{Lt}} \mu_Y \implies X \leq_{\mathrm{Lt}} Y$, if the diversity orders of X and Y are equal. Therefore, the generalized LT ordering of the Thorin measures of two GGCs with equal diversity order implies that the average SER performance at all SNR for the first one is better than that of the second at all SNR.

Laplace transform ordering of GGCs can also be obtained through the observation that any GGC with finite Thorin mass is a gamma variance mixture. This leads to the following Proposition:

Proposition : Let $X = AU_1$ and $Y = BU_2$ be two GGCs, where A and B are nonnegative RVs, independent of $U_j \sim Gamma(m_j, 1)$, $j = 1, 2$. Then $m_1 \leq m_2$ and

$A \leq_{\text{Lt}} B$ implies $X \leq_{\text{Lt}} Y$.

Proof. To begin with, observe that the LT of a Gamma RV monotonically decreases with the scale parameter, when the rate is unity. Therefore, if U_1 and U_2 are Gamma RVs with shape m_1 and m_2 respectively, and unit rate, then $U_1 \leq_{\text{Lt}} U_2 \iff m_1 \leq m_2$. Let A and B be non-negative RVs independent of U_1 and U_2, which satisfy $A \leq_{\text{Lt}} B$. Then

$$\mathbb{E}\left[\exp(-\rho A U_1)\right] = \mathbb{E}_A\left[\mathbb{E}_{U_1}\left[\exp(-\rho A U_1)\right]\right] \tag{5.71}$$

$$\geq \mathbb{E}_A\left[\mathbb{E}_{U_2}\left[\exp(-\rho A U_2)\right]\right] \tag{5.72}$$

$$\geq \mathbb{E}_B\left[\mathbb{E}_{U_2}\left[\exp(-\rho A U_2)\right]\right] , \tag{5.73}$$

where (5.72) is obtained from (5.71) by observing that $U_1 \leq_{\text{Lt}} U_2$, and applying (2.20) within the outer expectation in the right hand side of (5.71), since $\exp(-x)$ is a c.m. function. Similarly, (5.73) follows from (5.72), as $A \leq_{\text{Lt}} B$. □

While Proposition 5.12.1 connects the ordering of the Thorin measures and the LT ordering of the corresponding GGCs, Proposition 5.12.1 enables the LT ordering of pairs of GGC RVs for which the Thorin measure may not be available in closed form.

5.12.2 Comparison of GGC Channels with Equal Diversity Orders

It is well known that the AWGN (no fading) channel is a benchmark to the performance of any fading channel, with respect to symbol error rates and ergodic capacity. Intuitively, this is because the AWGN channel has infinite diversity order (since a Nakagami-m with $m \to \infty$ is an AWGN channel). However, when two GGC fading distributions with the same diversity order are to be compared, it is in fact observed that the Nakagami-m fading scenario is the best possible fading channel with respect to symbol error rates of 1-dimensional or 2-dimensional constellations

(which are convex functions) and the ergodic capacity (which is a concave function), as seen from the following Proposition.

Proposition : Let X be a GGC with Thorin mass D and $\mathbb{E}[X] = 1$. Then $U \leq_{\text{cx}} X$, where $U \sim Gamma(D, D)$.

Proof. According to Lemma 5.9.3, X can be written as AU, where U is Gamma distributed with shape D and rate D, and A is a nonnegative RV independent of Γ. Now, let g be a convex function. Then

$$\mathbb{E}_U[g(\mathbb{E}[A]\,U)] \leq \mathbb{E}[g(X)]\ , \tag{5.74}$$

as a consequence of Jensen's inequality. $\mathbb{E}[A] = 1$, because $\mathbb{E}[X]$ has been assumed to be 1. The proof is thus concluded. □

It is now possible to see from Proposition 5.12.2 that among all GGC fading channels with a given diversity order and unit average power, the Nakagami-m fading channel forms the benchmark channel with respect to convex performance metrics such as SERs of 1-dimensional or 2-dimensional constellations, and concave metrics such as the ergodic capacity. This is because the instantaneous channel power of a Nakagami-m channel is Gamma distributed with shape m and rate m. In other words, if the diversity order of a GGC is D, then the average SER of any 1-dimensional or 2-dimensional constellation is lower bounded by that of a Nakagami-m channel with $m = D$, and further, the ergodic capacity is upper bounded by that of the Nakagami-m channel with $m = D$.

5.12.3 Quantifying Asymptotic SER Performance Difference

While stochastic ordering techniques are capable of yielding GGC channel comparisons based on very general performance metrics without the need for closed form

expressions, an inherent limitation of this method is that it is not possible to quantify the exact performance gap of one channel with respect to the other. In the absence of a unified fading model, it is difficult to obtain the SNR gain of one fading distribution with respect to another, without explicitly obtaining the asymptotic SER expressions for both channels. In this account, the properties of GGCs make it possible to produce a closed form expression for the SNR gain of one GGC with respect to another, when comparing the SERs at high SNR. The SNR gain of one channel with respect to another is defined as the difference between the SNR (in dB) for the asymptotic SER of the two channels to be equal, and is denoted as $g_{db}(X_1, X_2)$. Mathematically,

$$g_{db}(X_1, X_2) = \lim_{\rho \to \infty} 10 \log_{10} \frac{\overline{P}_{\mathrm{e}}^{(1)}(\rho)}{\overline{P}_{\mathrm{e}}^{(2)}(\rho)}, \tag{5.75}$$

where $\overline{P}_{\mathrm{e}}^{(1)}(\cdot)$ and $\overline{P}_{\mathrm{e}}^{(2)}(\cdot)$ are the SERs of the channels with instantaneous channel power X_1 and X_2 respectively.

Proposition : Let X_1 and X_2 correspond to the instantaneous channel powers of two GGC channels satisfying $\int_1^\infty \log u \mu_j(\mathtt{d}u) < \infty$, and $\int_0^\infty \mu_j(\mathtt{d}u) = \alpha$, $j = 1, 2$. If MPSK or MQAM is employed, the SNR gain is given by

$$g_{db}(X_1, X_2) = \frac{10}{\alpha}(\log_{10} e) \int_0^\infty (\log u)(\mu_1(\mathtt{d}u) - \mu_2(\mathtt{d}u)), \tag{5.76}$$

where $\mu_j(\cdot)$ is the Thorin measure of X_j, $j = 1, 2$. Equivalently,

$$g_{db}(X_1, X_2) = \frac{10}{D} \log_{10} \frac{h_1(0+)}{h_2(0+)}, \tag{5.77}$$

where $h_j(\cdot)$ are obtained from the canonical GGC PDF (5.44) of $X_j, j = 1, 2$.

Proof. Relation (5.77) is proved first. Then (5.76) is obtained using the relation (5.45). Let ρ_1 and ρ_2 denote the average SNRs for channels X_1 and X_2 required to obtain the same asymptotic SER. In other words, $(G_1\rho_1)^{-\alpha} = (G_2\rho_2)^{-\alpha}$. Relation

(5.77) then follows by taking logarithms both sides, and substituting the array gain expressions from (5.66). □

5.13 Systems Involving Multiple GGC RVs

Many systems of practical interest involve combinations of multiple fading components. In what follows, the conditions under which the overall end-to-end instantaneous channel power is a GGC with finite Thorin mass is obtained. This permits the comparison of two GGC channel systems using Proposition 5.12.1. This facilitates the use of Proposition 5.12.3 to quantify the high SNR gain of one GGC channel with respect to another, since this propositions require finite Thorin mass.

5.13.1 Composite Fading Systems

Many fading models in wireless communications attempt to capture both the short term fading effect and the long term shadowing effect through a composite fading distribution. In such cases, the instantaneous channel power is modelled as a product of two RVs, one corresponding to the short-term effect, and the other corresponding to the long-term fading effect. In other words, the overall instantaneous channel power Z can be written as $Z = XY$. For such systems, Z is a GGC if X is a GGC and Y is HCM, as seen through an application of Lemma 5.9.4. The Thorin mass of Z is finite if the Thorin mass of either X or Y is finite, as seen from Lemma 5.9.3. As a result, common composite models such as the Rayleigh-lognormal model [85], the Nakagami-lognormal model [86], Weibull-Gamma composite model [87], Generalized-K [80], and the Gamma-shadowed generalized Nakagami fading model [75] have GGC instantaneous channel powers with finite Thorin mass.

It can then be concluded using Proposition 5.12.1 that given two different composite fading distributions finite Thorin mass, if the diversity order of the first one

is larger than the second, then to establish LT ordering of the composite distribution, it is sufficient to establish LT ordering of the mixing distribution. Further, if the Thorin measures of the two composite distributions satisfy $\int_0^\infty \log u \mu_j(\mathtt{d}u) < \infty$, then Proposition 5.12.3 can be used to quantify the high SNR gain of one system versus the other.

5.13.2 Diversity Combining Systems

Consider a single-input multiple-output diversity combining system with L receive antennas, and complete CSI at the receiver. It is now proved that under different assumptions on the instantaneous channel power of each branch, the end-to-end instantaneous channel power of combining schemes such as MRC, EGC and SC is a GGC.

Maximum Ratio Combining

In the case of MRC, the instantaneous end-to-end channel power $X_{\text{MRC}} = \sum_{i=1}^{L} X_i$ is a GGC whenever X_i is a GGC, since a sum of GGC RVs is a GGC. It is therefore straightforward to see that the average SER of DPSK is given by one-half $\phi(\rho)$ in (5.39), where the Thorin measure is the sum of Thorin measures of each component of the MRC system. The ergodic capacity and average SERs of other 2-dimensional modulations may however not be tractable in general, for example when the branches are independent but not identically distributed GGCs. In this case, it is possible to obtain a performance lower bound if the diversity order is known. To this end, from Proposition 5.9.3, if every component has finite diversity order, then the diversity order of the MRC system is finite, and is given by $D = \sum_{i=1}^{N} \int \mu_i(\mathtt{d}u)$. Now, since X_{MRC} is a GGC with Thorin mass $D < \infty$, using Proposition 5.12.2, it is seen that the performance of the MRC system is lower bounded by that of a Nakagami-m

channel with $m = D$, with respect to all convex metrics such as average SERs of 2-dimensional modulations and concave metrics such as the ergodic capacity.

Equal Gain Combining

Next, consider an EGC system, where $X_{\text{EGC}} = L^{-1}\left(\sum_{i=1}^{L}\sqrt{X_i}\right)^2$. Although it unknown whether X_{EGC} is a GGC whenever X_i is a GGC or even HCM, it can be shown that when X_i is a product of generalized Gamma distributed RVs, X_{EGC} is a GGC. Closed-form expressions for the average SER and ergodic capacity in this case are intractable for many distributions such as Nakagami-m and Hoyt, which are special cases of generalized Gamma distributions. In order to obtain a lower bound on the performance metrics, observe from Proposition 5.9.3 that X_{EGC} is a GGC with finite diversity order if every branch has finite diversity order. Furthermore, the diversity order $D = \sum_{i=1}^{N}\int \mu_i(\mathtt{d}u) < \infty$. Proposition 5.12.2 can then be used to bound the average SER and the ergodic capacity by that of a Nakagami-m channel.

Selection Combining

On the other hand, when SC is performed, $X_{\text{SC}} = \max_{i=1,\ldots,L} X_i$. Similar to the EGC case, it is unknown if X_i being a GGC or HCM results in a GGC distributed X_{SC}. However, it can be shown that when X_i is of the form Y/Z, where Y is a GGC, and Z is an exponential RV, then X_{SC} is a GGC. X_i can be written in this form in several scenarios of interest, for example, in an interference limited system with only one dominant Rayleigh faded interference source. In such cases, it is possible to show that X_{SC} is also a ratio of a GGC RV and an exponential RV. The diversity order of a SC system is the same as that of an MRC system, which is given by $D = \sum_{i=1}^{N}\int \mu_i(\mathtt{d}u)$. To ensure D to be finite, $\int \mu_i(\mathtt{d}u)$ must be finite for each $i = 1, \ldots, L$. In other words, the Thorin mass of each branch must be finite. Since

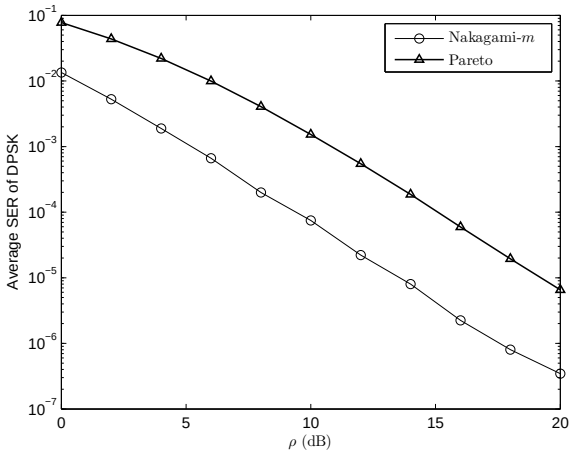

Figure 5.4: Average SER performance of 2-branch MRC with Pareto(1.25) distributed SNR compared with that of Nakagami-m SISO channel with $m = 2.5$

.

each branch instantaneous channel power is of the form Y/Z, using Proposition 5.9.3, it is seen that it is sufficient to ensure that the Thorin mass of Y is finite, since Z is exponentially distributed, and has Thorin mass of 1. Consequently, Proposition 5.12.2 can be employed to bound the performance of an EGC system using that of a Nakagami-m system.

5.14 Simulations

In this section some of the theoretical results are corroborated using Monte-Carlo simulations. In Figure 5.4, the performance of a 2-branch MRC system where the instantaneous channel powers are Pareto distributed with parameters $(1, 1, 1.25)$ is simulated, and compared with that of a SISO Nakagami-m channel with $m = 2.5$. The average powers of both the systems have been normalized to unity. This simulation demonstrates that the average SER of the MRC system with Pareto distributed branches, which are GGC's with finite Thorin mass is lower bounded by that of a Nakagami-m channel, as suggested in Section 5.13.2.

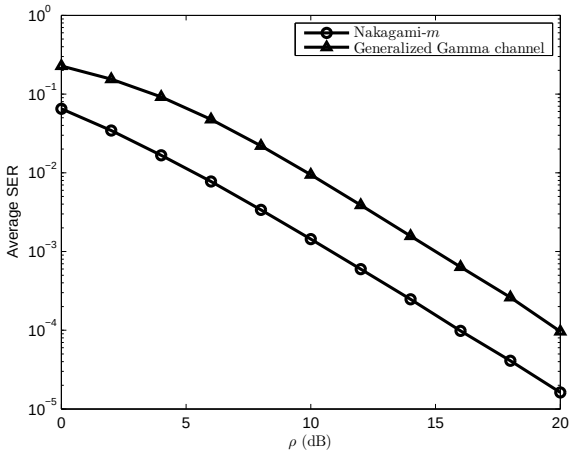

Figure 5.5: Comparison of DPSK performance of Nakagami-m ($m = 2$) channel and generalized gamma distributed channel Gamma$(1, 1)^{1/2}$ with equal avg. channel powers and diversity orders.

.

In Figure 5.5, the performance of a Nakagami-m channel with $m = 2$ is compared with that of a fading channel with generalized gamma distributed instantaneous channel power with parameters ($\epsilon = 2$, $r = 2$, $c_2 = 1$). The parameters of the two distributions have been chosen such that the diversity order is 2 for both cases. The high SNR gain in dB obtained from the simulation is found to be ≈ 1.7 dB. This agrees with the theoretically suggested value of 1.505 dB obtained from (5.77), with $h_1(0+) = 4$ and $h_2(0+) = 2$.

A simulation to provide an intuitive understanding of the structure of the Thorin measure and its effect on the average SER performance has been provided in Figure 5.7. To this end, two different GGCs $X = X_1 + X_2$ and $Y = Y_1 + Y_2$, where X_1, X_2 are independent and gamma distributed with parameters $(2, 1)$ and $(2, 2)$ respectively, and Y_1, Y_2 are independent and gamma distributed with parameters $(1, 1/2)$ and $(3, 3)$ respectively, are chosen. In this case, the densities of the Thorin measure are as depicted in Figure 5.6. The average SER for X and Y are obtained in figure 5.7. It is observed that the average SER of X is consistently less than that of Y at all $\rho \geq 0$.

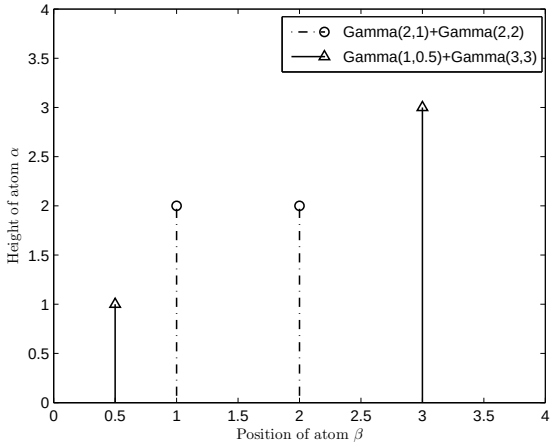

Figure 5.6: Density of Thorin measures for two GGCs. X =Gamma$(2,1)$+Gamma$(2,2)$; Y =Gamma$(1,0.5)$+Gamma$(3,3)$. Thorin mass $= 4$, $\mathbb{E}[X] = \mathbb{E}[Y] = 3$

.

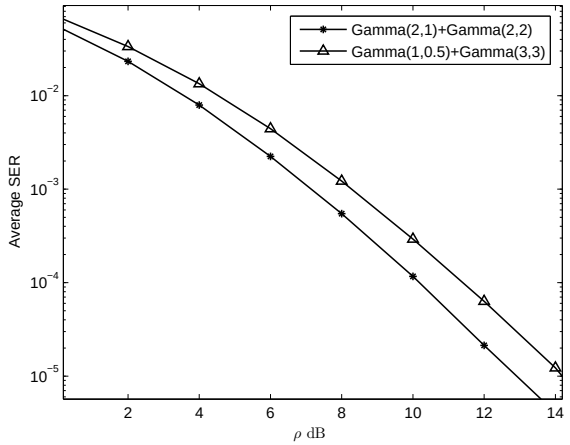

Figure 5.7: SER performance of DPSK over two GGCs which are sums of 2 gamma RVs with different supports of Thorin measures. The less spread apart μ is, the better the performance.

.

Chapter 6

CONCLUSIONS

In this Thesis, we studied four separate yet related problems (i) Stochastic ordering and its implications in physical layer wireless communications systems;(ii) A new canonical representation for the SER of an arbitrary constellation over an AWGN channel; (iii) The proposition of a new stochastic order to compare fading channels based on the ergodic capacity; and (iv) The unification of existing fading models under the umbrella of infinitely divisible distributions.

In Chapter 3 of this dissertation, we illustrate the power of stochastic orders such as the convex order and the LT order, which have never been used in physical layer communication/information theory, to relate and unify existing performance metrics such as ergodic capacity and error rate functions through their relationship with completely monotonic functions. We first identify that the instantaneous symbol error rate functions for various signaling constellations such as M-PSK and M-QAM are completely monotonic functions of the instantaneous SNR. Recognizing the importance of LT ordering of instantaneous SNR distributions, we identify parametric fading distributions such as Nakagami and Ricean, which are monotonic in the LoS parameters in the orders $\leq_{\mathrm{Lt}}$ and $\leq_{\mathrm{cx}}$. We also lay the groundwork to find the conditions for the preservation of inequalities satisfied by the averages of performance metrics of individual systems, when multiple such systems are combined, even when closed form expressions for such averages are not tractable. These include diversity combining schemes such as MRC, EGC and a variety of relay networks. In summary, this framework provides a novel approach to compare the performance of a vast range of systems on the basis of the analytical properties of the performance metric such as

monotonicity, convexity, or complete monotonicity, even in settings where closed-form expressions are not tractable.

In Chapter 4, the SER of an arbitrary constellation with reduced dimension $N^* \geq 2$ under AWGN is characterized as $P_{\mathrm{e}}(\rho) = \rho^p f_{\mathrm{cm}}(\rho)$, where $f_{\mathrm{cm}}(\cdot)$ is a c.m. function, and $p \geq N^*/2 - 1$. This representation of the SER is shown to apply to cases when the noise follows a compound Gaussian distribution. The expression for the SER obtained herein is useful in establishing that the SER is a c.m. function if the constellation has a reduced dimension of one or two. The complete monotonicity of SER for constellations with $N^* > 2$ is shown to depend on the differentiability properties of the representing function corresponding to $\rho^{-p} P_{\mathrm{e}}(\rho)$, which is a function of the constellation geometry and the prior probabilities. The exact relation between the constellation geometry and the complete monotonicity of the SER for constellations with $N^* > 2$ is left as an open problem. Complete monotonicity of the SER has applications in obtaining comparisons of averages of SERs over pairs of quasi-static fading channels, such as Nakagami-m, whose instantaneous SNRs are Laplace transform ordered. Such comparisons can be made even in cases where a closed-form expression for the average SER is not analytically tractable. In addition, a new stochastic ordering relation is introduced, which can be exploited to obtain comparisons of the average SER of an arbitrary multidimensional constellation over two different fading channels. Using such stochastic comparisons, the minimum possible SER of a constellation of a given dimension, between two fading channels may be ordered. This provides a way to compare the channels based on their ergodic capacities.

In Chapter 5, it is shown that the ergodic capacity order and its properties can be exploited to obtain comparisons of ergodic capacities of composite systems across two different fading channels whose instantaneous SNRs satisfy the ergodic capacity

order. For systems such as MRC and EGC which involve multiple instantaneous SNR RVs, we conclude that combining a better set of channels (in the ergodic capacity order) produces a system with a higher ergodic capacity. This conclusion is true for all systems whose end-to-end instantaneous SNR belongs to the $\mathcal{CTBF}_m$ set. For systems whose end-to-end SNR does not belong to $\mathcal{CTBF}_m$, component-wise ergodic capacity ordering of instantaneous SNR need not produce a system with a higher ergodic capacity. An example to illustrate this point is the AF relay for which the instantaneous SINR is Pareto distributed. An extension of the ergodic capacity order to MIMO systems is also proposed herein. The properties of the ergodic capacity order can be used to compare the capacity regions of systems such as the multi-user MAC in two different fading environments, for both the single and multiple antenna case.

Finally, in Chapter 6, we propose the class of ID RVs as a new unified class to model fading channels. This is because many popularly employed shadowing, multipath fading and composite shadowing/multipath fading distributions such as Rayleigh, Rician, SIRP, Nakagami-m, lognormal and Nakagami-lognormal distributions are observed to have ID instantaneous channel powers. Further, it is noted that almost every ID fading distribution is also a GGC fading distribution. This permits the leverage of the rich analytical properties of GGCs to help in asymptotic SER analysis of constellations such as MPSK. Further, GGC fading channels with finite diversity order are characterized as variance mixtures of Gamma RVs. This permits us to stochastically order GGC fading distributions in a relatively straightforward manner.

References

[1] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[2] M. Simon and M. Alouini, *Digital communication over fading channels*. Wiley-IEEE Press, 2000.

[3] M. Shaked and J. G. Shanthikumar, *Stochastic orders and their applications*, 1st ed. Springer, Oct. 1994.

[4] A. W. Marshall and I. Olkin, *Inequalities: Theory of majorization and its applications*. Academic Press, Jan. 1980.

[5] D. P. Palomar and Y. Jiang, *MIMO transceiver design via majorization theory*. Now Publishers Inc, Jun. 2007.

[6] F. Belzunce and M. Shaked, "Failure profiles of coherent systems," *Naval Research Logistics (NRL)*, vol. 51, no. 4, pp. 477–490, 2004.

[7] A. Müller and D. Stoyan, *Comparison methods for stochastic models and risks*. John Wiley & Sons Inc, 2002.

[8] J. Caballe and J. Esteban, "Stochastic dominance and absolute risk aversion," *Social Choice and Welfare*, vol. 28, pp. 89–110, 2007.

[9] J. Quirk and R. Saposnik, "Admissibility and measurable utility functions," *The Review of Economic Studies*, vol. 29, no. 2, pp. 140–146, 1962.

[10] S. Ross, *Stochastic processes*. Wiley New York, 1996.

[11] B. Blaszczyszyn and D. Yogeshwaran, "Directionally convex ordering of random measures, shot noise fields, and some applications to wireless communications," *J. Advances Applied Probability*, vol. 41, no. 3, pp. 623–646, 2009.

[12] M. Azizoglu, "Convexity properties in binary detection problems," *IEEE Trans. Inf. Theory*, vol. 42, pp. 1316–1321, Jul. 1996.

[13] Y. Lin and S. Phoong, "BER minimized OFDM systems with channel independent precoders," *IEEE Trans. Signal Process.*, vol. 51, pp. 2369–2380, Sep. 2003.

[14] S. Loyka, V. Kostina, and F. Gagnon, "Error rates of the maximum-likelihood detector for arbitrary constellations: convex/concave behavior and applications," *IEEE Trans. Inf. Theory*, vol. 56, pp. 1948–1960, Apr. 2010.

[15] D. Middleton, "Statistical-physical models of electromagnetic interference," *IEEE Trans. Electromag. Compat.*, no. 3, pp. 106–127, 1977.

[16] M. Taqqu, *Stable non-Gaussian random processes: stochastic models with infinite variance*. Chapman & Hall/CRC, 1994.

[17] R. Schilling, R. Song, and Z. Vondraček, *Bernstein Functions: Theory and Applications*. Walter de Gruyter, 2010.

[18] S. Koumandos and H. L. Pedersen, "Completely monotonic functions of positive order and asymptotic expansions of the logarithm of Barnes double gamma function and Euler's gamma function," *J. Math. Anal. Appl.*, vol. 355, no. 1, pp. 33–40, 2009.

[19] N. Lebedev and R. Silverman, *Special Functions and their Applications*. Dover, 1972.

[20] R. Bhatia, *Matrix Analysis*. Springer Verlag, 1997, vol. 169.

[21] D. Kressner, "Bivariate matrix functions," in *(online) http://sma.epfl.ch/ anchpcommon/publications/multivariate.pdf*, 2012.

[22] A. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. Now Publishers Inc, 2004, vol. 1.

[23] D. Widder, *The Laplace transform*. Princeton Univ., 1946.

[24] G. Ziegler, *Lectures on polytopes*. Springer, 1995, vol. 152.

[25] C. De Concini and C. Procesi, *Topics in hyperplane arrangements, polytopes and box-splines*. Springer Verlag, 2010.

[26] N. Levy, O. Somekh, S. Shamai, and O. Zeitouni, "On certain large random Hermitian Jacobi matrices with applications to wireless communications," *IEEE Trans. Inf. Theory*, vol. 55, no. 4, pp. 1534 –1554, Apr. 2009.

[27] A. Karnik and A. Kumar, "Performance analysis of the Bluetooth physical layer," in *IEEE. Conf. Pers. Commun.*, 2000, pp. 70 –74.

[28] D. Bai, S. Ghassemzadeh, R. Miller, and V. Tarokh, "Beam selection gain from Butler matrices," in *Proc. IEEE. Veh. Technol. Conf.*, 2008, pp. 1–5.

[29] D. Tse, R. Yates, and Z. Li, "Fading broadcast channels with state information at the receivers," in *46th Annu. Allerton Conf. Commun. Control Computing*. IEEE, 2009, pp. 221–227.

[30] M. Nesenbergs, "Error probability for multipath fading - the slow and flat idealization," *IEEE. Trans. Commun.Technol.*, vol. 15, no. 6, pp. 797–805, Dec. 1967.

[31] M. Pun, V. Koivunen, and H. Poor, "Performance analysis of joint opportunistic scheduling and receiver design for MIMO-SDMA downlink systems," *IEEE. Trans. Commun.*, no. 99, pp. 1 –13, 2010.

[32] Z. Wang and G. Giannakis, "A simple and general parametrization quantifying performance in fading channels," *IEEE. Trans. Commun.*, vol. 51, no. 8, pp. 1389–1398, 2003.

[33] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, 2002.

[34] C. Tepedelenlioglu and P. Gao, "Estimators of the Nakagami-m parameter and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 519–527, 2005.

[35] D. Hoesli, Y. Kim, and A. Lapidoth, "Monotonicity results for coherent MIMO Rician channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4334–4339, 2005.

[36] M. Hasna and M. Alouini, "Outage probability of multihop transmission over Nakagami fading channels," *IEEE Commun. Lett.*, vol. 7, no. 5, pp. 216–218, 2003.

[37] A. Rajan and C. Tepedelenlioglu, "Diversity combining over Rayleigh fading channels with symmetric alpha-stable noise," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2968–2976, 2010.

[38] A. Ribeiro, X. Cai, and G. Giannakis, "Symbol error probabilities for general cooperative links," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1264–1273, 2005.

[39] C. Tepedelenlioglu, A. Rajan, and Y. Zhang, "Applications of stochastic ordering to wireless communications," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 4249 –4257, Dec. 2011.

[40] E. Biglieri, "Advanced modulation formats for satellite communications," *Advanced Methods for Satellite and Deep Space Communications*, pp. 61–80, 1992.

[41] G. Taricco, E. Biglieri, and V. Castellani, "Applicability of four-dimensional modulations to digital satellites: a simulation study," in *Proc. IEEE. Global Telecommun. Conf.*, vol. 4, Nov. 1993, pp. 28–34.

[42] H. Bülow, "Polarization QAM modulation (POL-QAM) for coherent detection schemes," in *Proc. Optical Fiber Commun. and National Fiber Optic Eng. Conf.*, 2009.

[43] E. Agrell and M. Karlsson, "Power-efficient modulation formats in coherent transmission systems," *J. Lightw. Technol.*, vol. 27, no. 22, pp. 5115–5126, 2009.

[44] ——, "On the symbol error probability of regular polytopes," *IEEE Trans. Inf. Theory*, vol. 57, pp. 3411–3415, Jun. 2011.

[45] A. Rajan and C. Tepedelenlioglu, "On the complete monotonicity of symbol error rates," in *Proc. IEEE Int. Symp. Inf. Theory*, 2012, pp. 1430–1434.

[46] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univ Pr, 2004.

[47] D. Du and F. Hwang, *Computing in Euclidean geometry*. World Scientific, 1995, vol. 4.

[48] J. Weeks, *The Shape of Space*. CRC press, 2002.

[49] J. Craig, "A new, simple and exact result for calculating the probability of error for two-dimensional signal constellations," in *Proc. IEEE Military Commun. Conf.*, 1991, pp. 571–575.

[50] A. Rajan and C. Tepedelenlioglu, "A representation for the symbol error rate using completely monotone functions," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3922–3931, 2013.

[51] ——, "Ergodic capacity ordering of fading channels," in *Proc. IEEE Int. Symp. Inf. Theory*, 2012.

[52] M. Hasna and M.-S. Alouini, "End-to-end performance of transmission systems with relays over Rayleigh-fading channels," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 1126 – 1131, Nov. 2003.

[53] O. Waqar, D. McLernon, and M. Ghogho, "Exact evaluation of ergodic capacity for multihop variable-gain relay networks: A unified framework for generalized fading channels," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 4181 –4187, Oct. 2010.

[54] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley-India, 1999.

[55] W. Rhee and J. Cioffi, "Ergodic capacity of multi-antenna gaussian multiple-access channels," in *Proc. 35th Asilomar Conf. Signals Syst. Comput.*, vol. 1, 2001, pp. 507–512.

[56] K. Yao, M. K. Simon, and E. Bigiieri, "Unified theory on wireless communication fading statistics based on SIRP," in *Proc. IEEE Signal Process. Adv. Wireless Commun.* IEEE, 2004, pp. 135–139.

[57] F. Yilmaz and M.-S. Alouini, "A unified MGF-based capacity analysis of diversity combiners over generalized fading channels," *Communications, IEEE Transactions on*, vol. 60, no. 3, pp. 862–875, 2012.

[58] E. Stacy, "A generalization of the gamma distribution," *Ann. Math. Stat.*, vol. 33, no. 3, pp. 1187–1192, 1962.

[59] M. D. Yacoub, "The α-μ distribution: A physical fading model for the stacy distribution," *IEEE Trans. Veh. Technol.*, vol. 56, no. 1, pp. 27–34, 2007.

[60] A. Ghasemi and E. S. Sousa, "Capacity of fading channels under spectrum-sharing constraints," in *IEEE Int. Conf. Commun.*, vol. 10. IEEE, 2006, pp. 4373–4378.

[61] G. K. Karagiannidis, N. C. Sagias, and P. T. Mathiopoulos, "N*nakagami: a novel stochastic model for cascaded fading channels," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1453–1458, 2007.

[62] N. C. Sagias, G. K. Karagiannidis, P. T. Mathiopoulos, and T. A. Tsiftsis, "On the performance analysis of equal-gain diversity receivers over generalized gamma fading channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 10, pp. 2967–2975, 2006.

[63] F. W. Steutel and K. Van Harn, *Infinite Divisibility of Probability Distributions on the Real Line*. CRC Press, 2003.

[64] L. Bondesson, *Generalized Gamma Convolutions and Related Classes of Distributions and Densities*. Springer, 1992, vol. 76.

[65] ——, "A remarkable property of generalized gamma convolutions," *Probab. theory and related fields*, vol. 78, no. 3, pp. 321–333, 1988.

[66] ——, "Generalized gamma convolutions and complete monotonicity," *Probab. theory and related fields*, vol. 85, no. 2, pp. 181–194, 1990.

[67] L. F. James, B. Roynette, M. Yor *et al.*, "Generalized gamma convolutions, dirichlet means, thorin measures, with explicit examples," *Probab. Surv*, vol. 5, pp. 346–415, 2008.

[68] O. E. Barndorff-Nielsen and N. Shephard, "Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics," *J. Royal Stat. Soc.: Series B (Stat. Methodol.)*, vol. 63, no. 2, pp. 167–241, 2001.

[69] P. Carr, X. Jin, and D. B. Madan, "Optimal investment in derivative securities," *Finance and Stochastics*, vol. 5, no. 1, pp. 33–59, 2001.

[70] E. Figueroa-Lopez and C. Houdré, "Nonparametric estimation for lévy processes with a view towards mathematical finance," *arXiv preprint math/0412351*, 2004.

[71] D. E. Burmaster and A. M. Wilson, "An introduction to second-order random variables in human health risk assessments," *Human and Ecological Risk Assessment*, vol. 2, no. 4, pp. 892–919, 1996.

[72] M. Schroder, "Risk-neutral parameter shifts and derivatives pricing in discrete time," *J. Finance*, vol. 59, no. 5, pp. 2375–2402, 2004.

[73] Y. Zhang and C. Tepedelenlioglu, "Applications of Tauberian theorem for high-SNR analysis of performance over fading channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 296–304, 2012.

[74] O. Thorin, "On the infinite divisibility of the lognormal distribution," *Scandinavian Actuarial J.*, vol. 1977, no. 3, pp. 121–148, 1977.

[75] M.-S. Alouini and A. J. Goldsmith, "A unified approach for calculating error rates of linearly modulated signals over generalized fading channels," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1324–1334, 1999.

[76] B. Chytil, "The distribution of amplitude scintillation and the conversion of scintillation indices," *J. Atmos. Terr. Phys.*, vol. 29, no. 9, pp. 1175–1177, 1967.

[77] R. Hoyt, "Probability functions for the modulus and angle of the normal complex variate," *Bell Syst. Tech. J.*, vol. 26, no. 2, pp. 318–359, 1947.

[78] N. Hajri, N. Youssef, and M. Patzold, "A study on the statistical properties of double Hoyt fading channels," in *Proc. Int. Symp. Wireless Commun.* IEEE, 2009, pp. 201–205.

[79] N. C. Sagias, D. A. Zogas, G. K. Karagiannidis, and G. S. Tombras, "Channel capacity and second-order statistics in Weibull fading," *IEEE Commun. Lett.*, vol. 8, no. 6, pp. 377–379, 2004.

[80] P. M. Shankar, "Error rates in generalized shadowed fading channels," *Wireless Pers. Commun.*, vol. 28, no. 3, pp. 233–238, 2004.

[81] J. B. Andersen, "Statistical distributions in mobile communications using multiple scattering," in *Proc. 27th URSI General Assembly*, 2002.

[82] X. Hong, C.-X. Wang, and J. Thompson, "Interference modeling of cognitive radio networks," in *IEEE Veh. Technol. Conf.* IEEE, 2008, pp. 1851–1855.

[83] S. Fourati, "α-stable densities are hyperbolically completely monotone for $\alpha \in]0, 1/4] \cup [1/3, 1/2]$," *arXiv preprint math/1309.1045v2*, 2013.

[84] H. Bateman, A. Erdélyi, H. van Haeringen, and L. Kok, *Tables of integral transforms.* McGraw-Hill New York, 1954, vol. 1.

[85] A. Turkmani, "Probability of error for m-branch macroscopic selection diversity," *Proc. IEE Commun. Speech and Vision*, vol. 139, no. 1, pp. 71–84, 1992.

[86] M.-J. Ho and G. Stuber, "Co-channel interference of microcellular systems on shadowed Nakagami fading channels," in *IEEE Veh. Technol. Conf.* IEEE, 1993, pp. 568–571.

[87] P. S. Bithas, "Weibull-gamma composite distribution: alternative multipath/shadowing fading model," *Electron. lett.*, vol. 45, no. 14, pp. 749–751, 2009.